



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 160527

TO: Janet Epps-Ford
Location: REM-2C05&2C70
Art Unit: 1633
Wednesday, August 03, 2005

Case Serial Number: 09/489079

From: Deirdre Arnold
Location: Biotech-Chem Library
REM 1A64
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

Search Notes

Please feel free to contact me if you have any questions or would like to amend the search.

Thank you for using STIC services.

Regards,
Deirdre Arnold



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:05:16 ; Search time 111.289 Seconds
(without alignments)
1831.334 Million cell updates/sec

Title: US-09-489-079-24
Perfect score: 2052
Sequence: 1 MQKSPVNALEKNEQTLRA.....SVRFILMKMKIISYMKIAC 398

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1938.5	94.5	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	1124	54.8	1011	2 Q9BXX2	Q9bxx2 homo sapien
3	369	18.0	1715	2 Q9UPS8	Q9ups8 homo sapien
4	368.5	18.0	1710	2 Q9H1Q1	Q9h1q1 homo sapien
5	306	14.9	381	2 Q15694	Q15694 homo sapien
6	303	14.8	424	2 Q9NSI9	Q9nsi9 homo sapien
7	282.5	13.8	450	2 Q6PI49	Q6pi49 homo sapien
8	268.5	13.1	1080	2 Q6ZRL14	Q6zrl14 homo sapien
9	242.5	11.8	718	2 Q9HCD1	Q9hcd1 homo sapien
10	231.5	11.3	2006	2 Q7KSQ6	Q7ksq6 plasmodium
11	231.5	11.3	2019	2 Q7KSQ5	Q7ksq5 plasmodium
12	231.5	11.3	2055	2 Q8IHP3	Q8ihp3 plasmodium
13	231.5	11.3	2055	2 Q8T5C7	Q8t5c7 plasmodium
14	220.5	10.7	1043	2 Q69ZS2	Q69zs2 mus musculus
15	214.5	10.5	308	2 Q6S8V2	Q6s8v2 homo sapien
16	199	9.7	648	2 Q8WP26	Q8wp26 macaca fasc
17	197.5	9.6	1952	2 Q6BQ17	Q6bq17 debaryomyce
18	194	9.5	733	2 Q8I1D2	Q8i1d2 mus musculus
19	191.5	9.3	7210	2 Q9V7G8	Q9v7g8 drosophila
20	191.5	9.3	9270	2 Q8MLD9	Q8mld9 drosophila
21	183	8.9	5507	2 Q8IHN3	Q8ihn3 plasmodium
22	182.5	8.9	2954	2 Q42263	Q42263 xenopus lae
23	181	8.8	18519	2 Q8ISF6	Q8isf6 caenorhabdi
24	181	8.8	18534	2 Q8ISF7	Q8isf7 caenorhabdi
25	179.5	8.7	1313	2 Q9XIP6	Q9xip6 arabidopsis
26	177	8.6	470	2 Q7ROE3	Q7rqe3 plasmodium
27	177	8.6	2042	2 Q6BUQ9	Q6buq9 debaryomyce
28	176.5	8.6	1359	2 Q6PUA5	Q6puas tetrahymena
29	176	8.6	823	2 Q9H0H6	Q9h0h6 homo sapien
30	173	8.4	585	1 INVO HUMAN	P07476 homo sapien
31	173	8.4	795	2 Q9H2G7	Q9h2g7 homo sapien

32	173	8.4	976	1 SCP1_HUMAN	Q15431 homo sapien
33	173	8.4	1305	2 Q9FJ35	Q9fj35 arabidopsis
34	173	8.4	1938	2 Q9GJP9	Q9gjp9 cryptotlagus
35	173	8.4	1993	2 Q7PF91	Q7pf91 anopheles g
36	173	8.4	2016	2 Q7PF90	Q7pf90 anopheles g
37	173	8.4	10578	2 Q8ISFS	Q8isfs caenorhabdi
38	172.5	8.4	680	2 Q9V6S8	Q9v6s8 drosophila
39	171.5	8.4	924	2 Q15738	Q15738 dictyosteli
40	171.5	8.4	2663	1 CENE_HUMAN	Q02224 homo sapien
41	171	8.3	957	1 ERC2_MOUSE	Q6ph08 mus musculu
42	170	8.3	957	1 ERC2_HUMAN	Q15083 homo sapien
43	170	8.3	957	1 ERC2_RAT	Q8k3m6 rattus norv
44	169	8.2	1357	1 KTN1_HUMAN	Q86up2 homo sapien
45	169	8.2	1935	2 Q6IQX1	Q6iqx1 brachydanio

ALIGNMENTS

RESULT 1
Q9BXX3
ID Q9BXX3 PRELIMINARY; PRT: 1341 AA.
AC Q9BXX3;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Breast cancer antigen NY-BR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.;
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF26987; AAK27325.1; -;
DR GO; GO:0005634; C:nucleus; NAS
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR ANK repeat.
SQ SEQUENCE 1341 AA; 152776 MW; 33B5DD6FD3A59B CRC64;
Query Match 94.5%; Score 1938.5; DB 2; Length 1341;
Best Local Similarity 98.2%; Pred. No. 6.3e-89;
Matches 378; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY	1	MQKSPVNALEKNEQTLRADEILPSESKQDYBESSWDSSESLCETVSQKDVCLPKAAHQ	60
DB	592	MQKSPVNALEKNEQTLRADEILPSESKQDYBESSWDSSESLCETVSQKDVCLPKAAHQ	651
QY	61	KEIDKINGKLEGPVKDGLLKCNCMKVSIPTKALELMDMQTFKABPPKPSAFAPAIEH	120
DB	652	KEIDKINGKLEGPVKDGLLKCNCMKVSIPTKALELMDMQTFKABPPKPSAFAPAIEH	711
QY	121	QKSVNKALELNEQTLRADEILPSESKQDYBESSWDSSESLCETVSQKDVCLPKATHQK	180
DB	712	QKSVNKALELNEQTLRADEILPSESKQDYBESSWDSSESLCETVSQKDVCLPKATHQK	771
QY	181	EIDKINGKLEGPVNDGFLKACPRKMSVPTKALELMDMQTFKABPPKPSAFAPAIEH	240

```

Db 772 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQKAEPPKPSAFPAIEMQ 831
Qy 241 KSVPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQKE 300
Db 832 KSVPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQKE 891
Qy 301 MDKISGKLEDSLSKILDTIHSERARELQKDCHEOCTGMEQMKKKFCVLKKLSEAK 360
Db 892 MDKISGKLEDSLSKILDTVHSERARELQKDCHEOCTGMEQMKKKFCVLKKLSEAK 951
Qy 361 EIKSQLENQKVKWEQELCSVRFLT 385
Db 952 EIKSQLENQKVKWEQELCSVR-LTL 975

RESULT 2
Q9BXX2 ID Q9BXX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BXX2;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269088; AAK27326.1; -.
DR HSSP; O75832; 1UOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS00088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
FT NON TER 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DB438 CRC64;

Query Match 54.8%; Score 1124; DB 2; Length 1011;
Best Local Similarity 49.2%; Pred. No. 2.3e-48;
Matches 244; Conservative 57; Mismatches 89; Indels 106; Gaps 8;

Qy 1 MQSVPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQ 60
Db 385 MQKTVPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQ 444
Qy 61 KEIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQKAEPPKPSAFEP 116
Db 445 KEFTLSKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQKAEPPKPSAFEP 504
Qy 117 -----AIENQK-----SVPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQ 146
Db 505 KVSIPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQ 564
Qy 147 SKQKDYESSWDSLSLCTYSOKDVCVPKATHQKEIDKINGKLESPDNDGFLKAPCRMK 206
Db 565 SKQKDYESSWDSLSLCTYSOKDVCVPKATHQKEIDKINGKLESPDNDGFLKAPCRMK 624
Qy 207 VSIPTKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQKE 248
Db 625 ISLPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQKE 684
Qy 249 ELKNEOTL-----RADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQ 299
Db 685 TGOQERDIGIIRAPQDQTNKMPSELGRKEDTKSDSEIISVSDTQNYECLPEATYQ 744

```

```

Qy 300 EMDKISGKLE-----DSTLSKILDTIHSCE 325
Db 745 EIKTTNGKIESPKPSHPFAPATMQNSVPNKGLEWKNKQTLRADSTLSKILDALPSC 804
Qy 326 RARELQKDCHEOCTGMEQMKKKFCVLKKLSEAKIKSQLENQKVKWEQELCSVRP--- 382
Db 805 RGRELKXDCNQITAKMEQMKKKFCVLKKLSEAKIKSQLENQKVKWEQELCSVRPLN 864
Qy 383 -----LTLMKMKI 390
Db 865 QEEKRRNVLDILKEKI 880

RESULT 3
Q9UPS8 ID Q9UPS8 PRELIMINARY; PRT; 1715 AA.
AC Q9UPS8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE KIAA1074 protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99397452; PubMed=10470851;
RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:197-205(1999).
DR EMBL; AB028997; BAA83026.2; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS00088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT; 1.
KW ANK repeat.
FT NON TER 1
SQ SEQUENCE 1715 AA; 196962 MW; A82858F5F58E3203 CRC64;

Query Match 18.0%; Score 369; DB 2; Length 1715;
Best Local Similarity 30.4%; Pred. No. 2.3e-10;
Matches 130; Conservative 77; Mismatches 160; Indels 60; Gaps 15;

Qy 5 VPKNALELNEOTL-----RADEILPSSKQKDYESSWDSLSLCTYSOKDVC-C 53
Db 383 IESAPLQETNNDLTYVDEVHKNRSDMSALGLQBEDIESPMDSESIENPFQKVD 442
Qy 54 LPKAAHQETDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMD-----MQTFKAEPP 108
Db 443 LAGAADGKKNIGNEQAE-----DVFIYPSG-MSGSRNFQAKLEDTRNMGVPAHESP 496
Qy 109 EKSAPAPATPMQSVPNKALELNEOTLRADQMFPSKQKNVEENSWSESRETYSOKDVCVPKATHQ 167
Db 497 ERYLHLKPTTEMKDSVPNKAGMKDVQTSKAAEHDLVASEEQRSGSENNQPVSEER 556
Qy 168 QKDVCLPKATHQKEI-----DKINGKLESPDNDGFL-----KAPCRMKVSIPTK 212
Db 557 KK-----HRNMEVSNIIHDGATDDAEDDDDDGLIQKRSGETDQHPKRNKEYAS 611
Qy 213 --ALELMDMQ-TFKAEPPKPSAFPAIEMQSVPNKALELNEOTLR-----ADQMFPS-- 264
Db 612 GPALQMKVEKSTBKERTSKESVNSPVFGKASLLTGGLLQVDDDDSSUSEIDEDEGRPTKK 671

```



```
QY 265 -----ESKQKNVENSWDSESLRET--VSQKQVCPKATHQKEM----DKISGKLEBDSLS 315
Db 672 TSNEKNKVNQIQSDMDVDDLQSSSETASELPHSSYKNFMLLIEQLGMECKDSVSL 731
QY 316 KLIDTIHSCERARELQDCEQCTGMEQMKKFCVLKCKLSEAKELISOLENOKVKEQ 375
Db 732 KIQDAALSCERLELKNHCCELLTVKIKMEDKVNVLQRELSETKEIKSOLHEQKVEWR 791
QY 376 ELCSVRF 382
Db 792 ELCSLRF 798
RESULT 4
Q9H1Q1 PRELIMINARY; PRT; 1710 AA.
AC Q9H1Q1,
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BA145E8.1 (KIAA1074).
CN Name=ba145E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162272; CAC19649.1; -.
DR HSSP; P20749; 1K1B.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK_4.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS0088; ANK_REPEAT; 4.
DR PROSITE; PS0297; ANK_REPEAT; 1.
KW ANK repeat.
SQ SEQUENCE . 1710 AA; 196409 MW; 01CBF9BADB894872 CRC64;
Query Match 18.0%; Score 368.5; DB 2; Length 1710;
Best Local Similarity 30.4%; Pred. No. 2.4e-10;
Matches 130; Conservative 77; Mismatches 160; Indels 61; Gaps 15;
QY 5 VPKALELNEQTL-----RADEILPSESOKDYESSWDSSESLCETVSQKDV-C 53
Db 377 ISSAPLEQTNNDLTVDEVHKNNRSDMMASALGLQGEEDIESFPWDSSEISENFPQKYVDP 436
QY 54 LPKAAHQEKIDKINGKLEGSVPKDGILLKANCCKMKSIPTKALELMD-----MOTFKAEPP 108
Db 437 LAGADGKEKNIGNEQAE-----DVFYIPSC-MSGSRNFMKLEDTNRVGMVPAHMESE 490
QY 109 EKPSAFEPALIMQKSVNPALELNEQTLRADE-ILPSESOKDYESSWDSSESLCETVS 167
Db 491 ERYLHLKPTIMKDSVPNKGAGMKDQVTSKAAEHLEVASSEBQERESENNQPOVEER 550
QY 168 QKDVCLPKATHQKEI-----DKINGKLESPNDGFL---KAPCRMKVSIPTK----- 212
Db 551 KK-----HRRNEMEVSAITHDGNATDAEDDDGLIQKRKSGETHQPPRKENKEYAS 605
QY 213 ---ALELMDQ-TFKAAPPEKPSAFEPALIMQKSVNPALELNEQTLR---ADQMFPSP- 264
Db 606 SGPALQMKVEKTEKEKRTSKESVNSPFGKASLLTGGLLQVDDSSLSSEIDEDGRPTK 665
QY 265 -----ESKQKNVENSWDSESLRET--VSQKQVCPKATHQKEM----DKISGKLEBDSLS 314
Db 666 KTSNEKNKVNQIQSDMDVDDLQSSSETASELPHSSYKNFMLLIEQLGMECKDSVSL 725
QY 315 SKLIDTIHSCERARELQDCEQCTGMEQMKKFCVLKCKLSEAKELISOLENOKVKEQ 374
Db 726 LKIQDAALSCERLELKNHCCELLTVKIKMEDKVNVLQRELSETKEIKSOLHEQKVEWE 785
```

```
QY 375 QELCSVRF 382
Db 786 RELCSLRF 793
RESULT 5
Q15694 PRELIMINARY; PRT; 381 AA.
AC Q15694;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Protein immuno-reactive with anti-PTH polyclonal antibodies
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96187433; PubMed=8608414;
RA Kumar R., Haugen J.D., Wieben E.D., Londowski J.M., Cai Q.;
RT "Inhibitors of renal epithelial phosphate transport in tumor-induced
osteomalacia and uremia";
RL Proc. Assoc. Am. Physicians 107:296-305 (1995).
DR EMBL; U28831; AAB02177.1; -.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0008200; F:ion channel inhibitor activity; NAS.
DR GO; GO:0030319; P:di-, tri-valent inorganic anion homeostasis; NAS.
FT NON_TER 1
FT NON_TER 381
SQ SEQUENCE 381 AA; 42574 MW; BFE144AB3A21C6DD CRC64;
Query Match 14.9%; Score 306; DB 2; Length 381;
Best Local Similarity 27.7%; Pred. No. 6e-08;
Matches 109; Conservative 69; Mismatches 174; Indels 42; Gaps 9;
QY 19 RADEILPSESOKDYESS-----SWDSSESLCETVSQKDVCLPKAAHQEKIDKINGKLEGSVP 75
Db 4 RADSVLNIARGKKGKERRVSSHQKPAKATSDKENSVPNNMATETKQISGTV--SSQ 61
QY 76 KDGLKANCCKMKSIPTKALELMDMOTFKAAPPEKPSAFEPALIMQKSVNPALELNEQ 135
Db 62 KOPALKATSDKDSVNSNIPTEIKOQQSGTVSSQKPAWKATSVKDSVNSNATEIKDQ 121
QY 136 TLRADDEILPSESOKDYESSWDSSESLCETVSQKDVCLPKATHQKEIDKINGKLEESPD 195
Db 122 -----IRGTVSSQR-----QPALKATGDEKDSVNSNIAREIKDGEK-SGTV--SPQK 164
QY 196 DGFLKAPCRMKVSIPTKALELMDMOTFKAAPPEKPSAFEPALIMQKSVNPALELNEQ 255
Db 165 QSAQKVFIPKKVSLNLIATRIITGGWKSGETEYPENLPTLKATIENKNSVLNTATKMDVQT 224
QY 256 LRADQ-MPPSSCKQKNVEENSWDSSESLRETYSQK-----DVCVPKATHQKE----- 300
Db 225 STPEQDLEWASEGEQKRLSEYENNOFQVKNQIHSRDDLDITIQSSQTVSEGDGLCCNCK 284
QY 301 -----MDKISGKLEBDSLSKILDTIHSCEARELQKHQCEQCTGRMEQMKKFCVLK 355
Db 285 NVILLIDHEMKCKDCVHLLKIKTKFCLCKRLTELKDNHCEQLRVKIRKLNKASVLOKR 344
QY 356 LSEAKEIKSOLNOKVKNVQELCSVRFITLMKWK 389
Db 345 LSEKEIKSQLKHETLELEKELCSLRFATQBEKK 378
RESULT 6
Q9NS19 PRELIMINARY; PRT; 424 AA.
AC Q9NS19;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
```


RA	Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A., Ighii S., Yamamoto J., Isono Y., Kawai-Hio Y., Sakito K., Nishikawa T., Kimura K., Matsuo K., Nakamura Y., Sekine M., Kituchi H., Kanda K., Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuko Y., Nagai K., Isogai T.; Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AK128577; BAC87508.1; -;
DN	NON TER 1080 1080
FT	
SR	SEQUENCE 1080 AA; 126741 MW; 47BBAAA933AE595C CRC64;
Query Match	13.1%; Score 268.5; DB 2; Length 1080;
Best Local Similarity	28.0%; Pred.No. 1.4e-05;
Matches	93; Conservative 62; Mismatches 96; Indels 81; Gaps 13;
QY	87 KVSIPTKALELMDMQTFCAPPEKPSAFAPAIEIMOKSVNPKALKEQLTLEADILPSE 146 : : : : : : : : : : : : 64 KVSLKEELQQ--DMDRFNE-----IGMLK-VFQALESKEVKQL----- 99
Dd	
QY	147 SKQKVDESSWSLSCTETVSQKDVLCPATHQKEIDKINGKL-----EESFDNDGF 198 : : 100 --QKEVEEE-----RKHRRNMMEVSANIHDGATDDAEDDDDDL 138
Dd	
QY	199 L---KAPCRMKVSIPTK-----ALELMDMQ-TFKAPPKPSPSAFEPAIEIMOKSV 244 : : 139 IQRKSGETHQQPFRKENKEYASSGPALQMKEVKSTEKETRTSKESVNPFVKASLLIT 198
Dd	
QY	245 NKALELKNEOTLR---ADMPPS-----ESKKNNBENSWDSSLRET--VSQKDVCV 293 : : 199 GGLOVDSDSLSEDIDEGRPTKTTSNEKNKVQNQIQSMDDVDLDLTQSSTASEDCPL 258
Dd	
QY	294 KATHQKEM----DKISGLEPSTSISKILDTHSCERARELQKHCEOCTGRMOMKKKPC 350 :: :: 259 HSSVKNFMLLIEQLGMCKDSVSLKIQDAALSCELRLLEKKNHCCELLTVIKRMEDKV 318
Dd	
QY	351 VLKKLSEAEEKSOLENOKVMQECLSVRF 382
Dd	
Dd	319 VLORELSETKEIKSOLEHQBKWRELCSLRF 350
RESULT 9	
ID	QHCD1 PRELIMINARY; PRT; 718 AA.
AC	Q9HCD1;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE	KIAA1641 protein (Fragment).
GN	Name=KIAA1641;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20450683; PubMed=10997877;
RA	Nagase T., Kikuno R., Nakayama M., Hirosema M., Ohara O.;
RT	"Prediction of the coding sequences of unidentified human genes. xviii. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro." ;
RD	J Mol Biol. 1997 Nov 14; 273(5):1157-1166.
RL	DNA RES. 7:273-281(2000).
DR	EMBL; AB046861; BAB13467.1; -;
DN	NON TER 1 1
FT	
SR	SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;
Query Match	11.8%; Score 242.5; DB 2; Length 718;
Best Local Similarity	27.6%; Pred.No. 0.00018;
Matches	83; Conservative 57; Mismatches 124; Indels 37; Gaps 8;
QY	109 EKPSAFEPAIEIMOKSVNPKALENKQTLTLEADILPSESKDYESSWDSSELCTVSQ 168 : : : : : : : : : : : : 13 QKPALKATSOKKQSVNIPTIEDQGQ-----SGTVSSSQK---QPAWKATSV-----K 58
Dd	
QY	169 KDVCPLPATHQKEIDKINGKLEESPNDNFLKAPCRMKVSIPTKALELMDMTQTFKAEP 228

Db	59	KDSVSNATEIKD-QGIRGTV--SPOKQSAQKVFKKCVSLNIAIRITGGHWSGTPE	115
Qy	229	KPSAFPAIEMQKSVNKALELKNQOTLRADQ--MPPSESKQKNVENSWDSLESRETYS	286
Db	116	NLPTLKATIENKNSVLNTATKMKDVQTSITPEQDLEWASGEQKRLSEYENNQPVKNQIH	175
Qy	287	QK-----DVCVPKATHQK-----MDKISGKLEDSTSLSKLDTTHSCERAR	328
Db	176	SRDLDDIIQSSQTVDSDGSLCCNCKNVILLIDQHEMKCKDCVHLKLIKNTFCLMKRLI	235
Qy	329	ELQKDHCCECTGMEQMKKFCVLKKLSEAKIKSQLENQKQKWEQELCSVRFLTLMKM	388
Db	236	KLKDNHCEQRVKIRKDKNKASVLRGRISKEBIKSQLKHEILELEKELCSLRAFAQEK	295
Qy	389	K 389	
Db	296	K 296	
RESULT 10			
Q7K5Q6		PRELIMINARY; PRT; 2006 AA.	
ID	Q7K5Q6		
AC	Q7K5Q6		
DT	05-JUL-2004	(T-EMBLrel. 27, Created)	
DT	05-JUL-2004	(T-EMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(T-EMBLrel. 27, Last annotation update)	
DB	Erythrocyte binding protein 3.		
GN	Names=maebi;		
OS	Plasmodium falciparum.		
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
OX	NCBI_Taxid=5833;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=22072559; PubMed=12076768; DOI=10.1016/S0166-6851(02)00067-1;		
RA	Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;		
RT	"Plasmodium falciparum MAEBL is a unique member of the ebl family.";		
RL	Mol. Biochem. Parasitol. 122:35-44 (2002).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=12077637; PubMed=12082132;		
RA	Michon P., Stevens J.R., Kaneko O., Adams J.H.;		
RT	"Evolutionary relationships of conserved cysteine-rich motifs in		
RL	adhesive molecules of malaria parasites.";		
RN	Mol. Biol. Evol. 19:1128-1142 (2002).		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RA	Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;		
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY042084; AAQ73469.1; -		
DR	InterPro; IPR008602; Duffy binding.		
DR	Pfam; PF05424; Duffy binding; 1.		
SQ	SEQUENCE 2006 AA; 237620 MW; EE75EF42E2BF767D CRC64;		
Query Match	11.3%;	Score 231.5; DB 2; Length 2006;	
Best Local Similarity	26.5%;	Pred. No. 0.002;	
Matches 109; Conservative	65;	Mismatches 166; Indels 71; Gaps 16;	
Qy	8	KALELKN-EQTLRADTELPSKQKQYE-----ESSWDSLSLCTVTSQK---DVCLPKAA	58
Db	1262	KADELKKAEKKKADELKKSEKKKADELKKKAEKKKADELKKKAEKKKADELKKKAE	1321
Qy	59	HQKEIDKINGKLKESPVYKQGLLKANCQKMSIPTKALELMDMTQFKAAPPEKPSAFE---	115
Db	1322	EKKKADEVK-KAEKKKADELKKSEKKKADELKKKAEKKKADELKKKAEKKKADELKKK	1380
Qy	116	-----PAIEMQKSVPNK-----ALELKNQOTLRADTELPSKQKQYEESWSWSESILCET	165
Db	1381	KAEKKKADELKKKAEKKKADELKKKAEKKKADELKKKAEKKKADELKKKAEKKKADELK	1436
Qy	166	VSQKDVCLPKATHQKEIDKINGKLKESPDNDGFLKAPCRMKVSIPTKALELMDMTQFKAA	225
Db	1437	-KKKAENLKKAEKKKADELKKKAEKKKADELKKKAEKKKADELKKKAEKKKADELKKKAE	1486

```

Qy 226 PPEKPSAFEPALIMOKS-----VPMKALELKN-EQTLRADOMFPSE-----SKQKNVE 272
Db 1487 EKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKVE 1546
Qy 273 ENSWDSLSRETYSQDVCVPKATHQKEMDKISGKLEDSTSLSKILDTIHSCERARELOK 332
Db 1547 QKKREBE--RRNMALRRAEILKQIEKKRIEVMKLYEEKKM-----KAEQLKK 1593
Qy 333 DHCEQCTG-----KMEQMKKFCVLKKLSE-----AKEIKSQLENQKVWQ 375
Db 1594 EEBEIKAEQLKKEEBEKKVQVQLKKKEEBEKKKAELKKEEBEENKTKAEQ 1644

RESULT 11
Q7K5Q5
ID Q7K5Q5 PRELIMINARY; PRT; 2019 AA.
AC Q7K5Q5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythrocyte binding protein 2.
GN Name=maebi;
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072553; PubMed=12076769; DOI=10.1016/S0166-6851(02)00067-1;
RA Blair P.L., Kappe S.H., Maciel J.E., Balu B., Adams J.H.;
RT "Plasmodium falciparum MAEBL is a unique member of the ebl family.";
RM Mol. Biochem. Parasitol. 122:35-44 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22077637; PubMed=12082132;
RA Michon P., Stevens J.R., Kaneko O., Adams J.H.;
RT "Evolutionary relationships of conserved cysteine-rich motifs in adhesive molecules of malaria parasites.";
RM Mol. Biol. Evol. 19:1128-1142 (2002).
RN [3]
RP SEQUENCE FROM N.A.
RA Blair B.L., Kappe S.H.I., Maciel J.E., Balu D.J., Adams B.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY042084; AA073468.1; -.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 2019 AA; 239410 MW; A626F2C684C08785 CRC64;

Query Match 11.3%; Score 231.5; DB 2; Length 2019;
Best Local Similarity 26.5%; Pred. No. 0.002;
Matches 109; Conservative 65; Mismatches 166; Indels 71; Gaps 16;

Qy 8 KALELKN-EQTLRADEILPSESQKQY-----ESSWDSLSLCTVTSQK---DVCLPKAA 58
Db 1262 KADELKKAEEKKKADELKKEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1321
Qy 59 HQEIDKINGKLGS PVKGLKNCGMKVISIPTKALELMDMTFFKAEPPEKPSAFE--- 115
Db 1322 EKKKADEVK-KAEKKKADELKKEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1380
Qy 116 -----PAIEMQKSVPNK-----ALELKNQTLRADEILPSESQKQYESSWDSLSLCT 165
Db 1381 KAEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1436
Qy 166 VSKQDVCLPKATHQKIDKINGKLERSPDNGFLKAPCRMKVISIPTKALELMDMTFFKAE 225
Db 1437 -KKAENLKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1486
Qy 226 PPEKPSAFEPALIMOKS-----VPMKALELKN-EQTLRADOMFPSE-----SKQKNVE 272
Db 1487 EKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1546
Qy 273 ENSWDSLSRETYSQDVCVPKATHQKEMDKISGKLEDSTSLSKILDTIHSCERARELOK 332
Db 1547 QKKREBE--RRNMALRRAEILKQIEKKRIEVMKLYEEKKM-----KAEQLKK 1593
Qy 333 DHCEQCTG-----KMEQMKKFCVLKKLSE-----AKEIKSQLENQKVWQ 375

```

```

Db 1547 QKKREBE--RRNMALRRAEILKQIEKKRIEVMKLYEEKKM-----KAEQLKK 1593
Qy 333 DHCEQCTG-----KMEQMKKFCVLKKLSE-----AKEIKSQLENQKVWQ 375
Db 1594 EEBEIKAEQLKKEEBEKKVQVQLKKKEEBEKKKAELKKEEBEENKTKAEQ 1644

RESULT 12
Q8IHP3
ID Q8IHP3 PRELIMINARY; PRT; 2055 AA.
AC Q8IHP3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MAEBL, putative.
GN ORFNames=PF11_0486;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.S., Nene V., Shalton J., Suh B., Peterson J., Angiuoli S.,
RA Partea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M., Fairall A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL; AE014843; AAN36066.1; -.
DR HSP; P04268; IIC2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0009405; P:pathogenesis; IEA.
DR InterPro; IPR008602; Duffy binding.
DR Pfam; PF05424; Duffy binding; 1.
SQ SEQUENCE 2055 AA; 243211 MW; 783F48BB7286B1F3 CRC64;

Query Match 11.3%; Score 231.5; DB 2; Length 2055;
Best Local Similarity 26.5%; Pred. No. 0.002;
Matches 109; Conservative 65; Mismatches 166; Indels 71; Gaps 16;

Qy 8 KALELKN-EQTLRADEILPSESQKQY-----ESSWDSLSLCTVTSQK---DVCLPKAA 58
Db 1262 KADELKKAEEKKKADELKKEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1321
Qy 59 HQEIDKINGKLGS PVKGLKNCGMKVISIPTKALELMDMTFFKAEPPEKPSAFE--- 115
Db 1322 EKKKADEVK-KAEKKKADELKKEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1380
Qy 116 -----PAIEMQKSVPNK-----ALELKNQTLRADEILPSESQKQYESSWDSLSLCT 165
Db 1381 KAEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1436
Qy 166 VSKQDVCLPKATHQKIDKINGKLERSPDNGFLKAPCRMKVISIPTKALELMDMTFFKAE 225
Db 1437 -KKAENLKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1486
Qy 226 PPEKPSAFEPALIMOKS-----VPMKALELKN-EQTLRADOMFPSE-----SKQKNVE 272
Db 1487 EKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKADELKKAEEKKKAE 1546
Qy 273 ENSWDSLSRETYSQDVCVPKATHQKEMDKISGKLEDSTSLSKILDTIHSCERARELOK 332
Db 1547 QKKREBE--RRNMALRRAEILKQIEKKRIEVMKLYEEKKM-----KAEQLKK 1593
Qy 333 DHCEQCTG-----KMEQMKKFCVLKKLSE-----AKEIKSQLENQKVWQ 375

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:24:58 ; Search time 6025.64 Seconds
(without alignments)
3200.517 Million cell updates/sec

Title: US-09-489-079-24

Perfect score: 2052

Sequence: 1 MQKSVPNKALELNEQTLRA.....SVRFLLTMKMKIISYMKIAC 398

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09489079/runat_01082005_141250_13297/app_query.fasta_1.1635
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079@cgn 1 1 8858 @runat_01082005_141250_13297 -NCPU=6 -ICPU=3
-NO MMAP -LARGSEQURY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.to.*
11: gb.ats.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2052	100.0	2683 6 BD271311	BD271311 Reagents
2	2052	100.0	2683 6 BD271312	BD271312 Reagents
3	2023	98.6	2232 6 AR283468	AR283468 Sequence
4	2023	98.6	2232 6 AR454017	AR454017 Sequence

5	2023	98.6	2232	6	AX303171	Sequence
6	2023	98.6	3681	6	AR283451	Sequence
7	2023	98.6	3681	6	AR344219	Sequence
8	2023	98.6	3681	6	AR351420	Sequence
9	2023	98.6	3681	6	AR454000	Sequence
10	2023	98.6	3681	6	AR561588	Sequence
11	2023	98.6	3681	6	AX282970	Sequence
12	2023	98.6	3681	6	AX303143	Sequence
13	1965.5	95.8	3288	6	AR283467	Sequence
14	1965.5	95.8	3288	6	AR454016	Sequence
15	1965.5	95.8	3288	6	AX303170	Sequence
16	1965.5	95.8	3865	6	AR283457	Sequence
17	1965.5	95.8	3865	6	AR351426	Sequence
18	1965.5	95.8	3865	6	AR454006	Sequence
19	1965.5	95.8	3865	6	AR561594	Sequence
20	1965.5	95.8	3865	6	AX303154	Sequence
21	1963.5	95.7	4458	6	AX829110	Sequence
22	1963.5	95.7	4458	9	AP269087	Sequence
23	1948	94.9	2307	6	AR283456	Sequence
24	1948	94.9	2307	6	AR344224	Sequence
25	1948	94.9	2307	6	AR351425	Sequence
26	1948	94.9	2307	6	AR454005	Sequence
27	1948	94.9	2307	6	AR561593	Sequence
28	1948	94.9	2307	6	AX282975	Sequence
29	1948	94.9	2307	6	AX303148	Sequence
30	1661.5	81.0	1337	6	AR283455	Sequence
31	1661.5	81.0	1337	6	AR344223	Sequence
32	1661.5	81.0	1337	6	AR351424	Sequence
33	1661.5	81.0	1337	6	AR454004	Sequence
34	1661.5	81.0	1337	6	AR561592	Sequence
35	1661.5	81.0	1337	6	AX282974	Sequence
36	1661.5	81.0	1337	6	AX303147	Sequence
37	1594.5	77.7	2030	6	AX367056	Sequence
38	1486	72.4	1448	6	CQ490684	Sequence
39	1486	72.4	1448	6	CQ495497	Sequence
40	1486	72.4	1448	6	CQ496522	Sequence
41	1419	69.2	1729	6	AR283454	Sequence
42	1419	69.2	1729	6	AR344222	Sequence
43	1419	69.2	1729	6	AR351423	Sequence
44	1419	69.2	1729	6	AR454003	Sequence
45	1419	69.2	1729	6	AR561591	Sequence

ALIGNMENTS

RESULT 1	BD271311	Reagents and methods useful for detecting diseases of the breast.	2683 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD271311					
DEFINITION	BD271311					
ACCESSION	BD271311.1	GI:33081079				
VERSION	BD271311.1	GI:33081079				
KEYWORDS	JP 2002540761-A/8.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 2683)					
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	Medel P.A.B., Cohen M., Colpitts T.L., Friedman P.N., Gordon J., Granados E.N., Hodges S.C. Klass M.R., Kratochvil J.D., Russell J.C. and Stroupe S.D.					
	Reagents and methods useful for detecting diseases of the breast					
TITLE	Patent: JP 2002540761-A 03-DEC-2002;					
JOURNAL	ABBOTT LABORATORIES					
COMMENT	OS Homo sapiens (human)					
	PN JP 2002540761-A/8					
	PD 03-DEC-2002					
	PF 21-JAN-2000 JP 2000594836					
	PR 21-JAN-1999 US 09/234716					
	PI PATRICIA A BILLING MEDEL, MAURICE COHEN, TRACEY L COLPITTS, PAULA					
	PI N FRIEDMAN,					
	PI JULIAN GORDON, EDWARD N GRANADOS, STEVEN C HODGES, MICHAEL R PI					
	KLASS,					


```
FT
FEATURES
  source
    /organism='Homo sapiens (human)'.
    Location/Qualifiers
      1..2683
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
  Pred. No.: 2,8e-165 Length: 2683
  Score: 2052.00 Matches: 398
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-489-079-24 (1-398) x BD271312 (1-2683)
Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 57 ATGC AAAAGTCTGTTCCAAATAAAGCCTTGGAAATTGAAGAATGAACAAACATTGAGAGCA 116
Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 117 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTTCT 176
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 177 GAGAGTCTCTGTGAGACTGTTTCACAGAGGATGTTGTATCCACCAAGCCTTGAATTTGAGACATG 236
Qy 61 LysGluLeuLeuAspLysAlaLeuGluLysGlnLysSerProValLysAspGlyLeuLeu 80
Db 237 AAAGAAATAGATAAATAAATGAAGAAATTTAGAAAGGCTCTCCTGTTAAAGATGGTCTTCTG 296
Qy 81 LysAlaAsnCysGlyMetLysValSerLeuProLysAlaLeuGluLeuMetAspMet 100
Db 297 AAGGCTAACTCGGAATGAAAGTTTCTATTTCACACTAAAGCCTTGAATTTGAGACATG 356
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 357 CAACCTTTCAAGCAGAGCCTCCCGAAGAGCCATCTGCTTCGACCTGCCATTGAATG 416
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAsp 140
Db 417 CAAGAGTCTGTTCCAAATAAAGCCTTGGAAATTGAAGAATGAACAAACATTGAGAGCAGAT 476
Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 477 GAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTTCTGAG 536
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 537 AGTCTCTGTGAGACTGTTTCACAGAGGATGTTGTATCCCAAGCCTTACCAAGCTTACATCAAAA 596
Qy 181 GluLeuAspLysAlaAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 597 GAAATAGATAAATAAATGAAGAAATTTAGAAAGTCTCCTGATAATGATGGTTTCTGAAG 656
Qy 201 AlaProCysArgMetLysValSerLeuProLysAlaLeuGluLeuMetAspMetGln 220
Db 657 GCTCCCTCGAGAATGAAGTTTCTATTTCACACTAAAGCCTTGAATTTGATGAGACATGCA 716
Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 717 ACTTTCAAGCAGAGCCTCCCGAAGGCCATCTGCCCTTCGACCTGCCATTGAATGCA 776
Qy 241 LysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGln 260
Db 777 AGTCTGTTCCAAATAAAGCCTTGGAAATTGAAGAATGAACAAACATTGAGACAGATCAG 836
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 837 ATGTTCCCTTCAGAAATCAAAACAAAGAACGTTGAAGAAATTTCTTTGGGATTTCTGAGAGT 896
```

```
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 897 CTCGGTCGAGACTGTTTCACAGAGGATGTTGTGTGTACCCAGGCTTACATCAAAAAGAA 956
Qy 301 MetAspLysLysSerGlyLysLeuGluAspSerThrSerLeuSerLysLysLysLeuAspThr 320
Db 957 ATGGATAAAATAAGTGGAAATTTAGAAGATTTCAACTAGCCCTATCAAAATCTTGATACA 1016
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1017 ATTCAATCTTGTGAAGAGCAAGGAACTTCAAAAAGATCACTGTGAACAATGTACAGCA 1076
Qy 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 1077 AAAATGGAAACAAATCAAAAAGAAAGTTTGTGTACTGTGAAAAAGAAACTGTCAAGAACAAA 1136
Qy 361 GluLeuLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLysSerVal 380
Db 1137 GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTTCGCAAGTGTG 1196
Qy 381 ArgPheLeuThrLeuMetLysMetLysLysLysLysLysLysLysLysLysLysLys 398
Db 1197 AGTCTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 1250
RESULT 3
AR283468 2232 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 491 from patent US 6528054.
ACCESSION AR283468
VERSION AR283468.1 GI:29720295
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2232)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and Hepler,W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6528054-A 491 04-MAR-2003;
FEATURES Location/Qualifiers
  source 1..2232
  /organism="unknown"
  /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
  Pred. No.: 6,66e-163 Length: 2232
  Score: 2023.00 Matches: 393
  Percent Similarity: 98.99% Conservative: 1
  Best Local Similarity: 98.74% Mismatches: 4
  Query Match: 98.59% Indels: 0
  DB: 6 Gaps: 0
US-09-489-079-24 (1-398) x AR283468 (1-2232)
Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAla 20
Db 1036 ATGC AAAAGTCTGTTCCCAATAAAGCCTTGGAAATTGAAGAATGAACAAACATTGAGAGCA 1095
Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1096 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTTCT 1155
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1156 GAGAGTCTCTGTGAGACTGTTTCACAGAGGATGTTGTGTATCCCAAGCCTTCRCATCAA 1215
Qy 61 LysGluLeuAspLysAlaAsnGlyLysLeuGluGluSerProValLysAspGlyLeuLeu 80
Db 1216 AAAGAAATAGATAAATAAATGAAGAAATTTAGAAAGGCTCTCCTGTTAAAGATGGTCTTCTG 1275
```

81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
1276 AAGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATG 1335
101 GlnThrPheLysAlaGluProGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
1336 CAAACTTTCAAAGCAGACGCTCCCGAAGAGCCATCTGCCCTCGAGCCTGCCATTGAATG 1395
121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
1396 CAAAGTCTCTTCCAAATAAGCCTTGAATTTGAAGAAATGAACAAACATTTGAGAGCAGAT 1455
141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
1456 GAGATCTCCATCAGAAATCAAAACAAAGAGCTATGAGAAAGTTCTTGGGATTTCTGAG 1515
161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
1516 AGTCTCTGTGAGACTGTTTCCAGAGAGGATGTGTGTTTACCACCAAGGCTRCRCATCAAAA 1575
181 GluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
1576 GAAATAGATAAAATAAATGGAATTTGAAGAAATGAACAAACATTTGAGAGCAGAT 1635
201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
1636 GCTCCCTGCAAGATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATGCAA 1695
221 ThrPheLysAlaGluProGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
1696 ACTTTCAAAGCAGACGCTCCCGAAGAGCCATCTGCCCTTCGAGCCTGCCATTGAATGCAA 1755
241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
1756 AAGTCTCTTCCAAATAAGCCTTGGAAATGAAGAAATGAACAAACATTTGAGAGCAGATCAG 1815
261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
1816 ATGTTCCCTTCAGNATCAAAACAAAGAAAGTTGAGAAATTTCTTGGGATTTCTGAGAGT 1875
281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
1876 CTCGTCGAGACTGTTTCCAGAGAGGATGTGTGTACCCCAAGGCTACACATCAAAAGAA 1935
301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
1936 ATGGATAAAATAAGTGGAAATTTAGAAAGATTCAACTAGCCTATCAAAATTTCTGGATACA 1995
321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
1996 GTTCATTCTTGTGAAAGAGCAGGGAACCTTCAAAAGATCACTGTGAACCAACGTACAGGA 2055
341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
2056 AAAATGCAACAAATGAAAGAAAGTTTGTGTACTGTAAGAAAGAACTGTGAGAGCAAAA 2115
361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnLeuLysCysSerVal 380
2116 GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAAGAGCTCTGCAGGTG 2175
381 ArgPheLeuThrLeuMetLysMetLysIleSerTyrMetLysIleAlaCys 398
2176 AGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 2229

RESULT 4
AR454017
LOCUS AR454017 2232 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 491 from patent US 6680197.
ACCESSION AR454017
VERSION AR454017.1 GI:42686807
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
1 (bases 1 to 2232)
Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L.,
Hepler,W.T. and Henderson,R.A.
Compositions and Methods for the therapy and diagnosis of breast
cancer
Patent: US 6680197-A 491 20-JAN-2004;
JOURNAL Location/Qualifiers
FEATURES
source 1..2232
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 6.66e-163 Length: 2232
Pred. No.: 2033.00 Matches: 393
Score: 2033.00
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 6 Gaps: 0
US-09-489-079-24 (1-398) x AR454017 (1-2232)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
DB 1036 ATGCNAAAGTCTCTCCCAATAAAGCCTTGAATTTGAATAATGAACAAACATTTGAGAGCA 1095
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
DB 1096 GATGAGATATCTCCCATCAGAAATCAAAACAAAGAGACTATGAAGAAAGTTCTTGGGATTTCT 1155
QY 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
DB 1156 GAGAGTCTCTGTGAGACTGTTTCCAGAGAGATGTGTGTTTACCACAGGCTRCRCATCAA 1215
QY 61 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
DB 1216 AAAGAAATAGATAAAATAAATGAAGAAATTTAGAAAGGCTCTCTGTTAAAGATGGTCTTCTG 1275
QY 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
DB 1276 AAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATG 1335
QY 101 GlnThrPheLysAlaGluProGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
DB 1336 CAAACTTTCAAAGCAGACGCTCCCGAAGAGCCATCTGCCCTTCGAGCCTGCCATTGAATG 1395
QY 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
DB 1396 CAAAGTCTCTTCCAAATAAAGCCTTGAATTTGAAGAAATGAACAAACATTTGAGAGCAGAT 1455
QY 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
DB 1456 GAGATATCTCCCATCAGAAATCAAAACAAAGAGACTATGAAGAAAGTTCTTGGGATTTCTGAG 1515
QY 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
DB 1516 AGTCTCTGTGAGACTGTTTCCAGAGAGATGTGTGTTTACCACCAAGGCTRCRCATCAAAA 1575
QY 181 GluIleAspLysIleAsnGlyLysLeuGluGlySerProAspAsnAspGlyPheLeuLys 200
DB 1576 GAAATAGATAAAATAAATGGAATTTGAAGAAATTTAGAAAGGCTCTCTGATAATGATGGTCTTCTGAAG 1635
QY 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
DB 1636 GCTCCCTGCAAGATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATGCAA 1695
QY 221 ThrPheLysAlaGluProGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
DB 1696 ACTTTCAAAGCAGACGCTCCCGAAGAGCCATCTGCCCTTCGAGCCTGCCATTGAATGCAA 1755
QY 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260

Db 1756 AAGTCTGTCCAAATAAAGCCTTGGAAATTGAAGAAATGAACAAACATTTGAGACGATCAG 1815
Qy 261 MetPheProSerGluSerLysGlnLysAenValGluGluAenSerTrpAspSerGluSer 280
Db 1816 ATGTTCCCTTCAGAAATCAAAACAAAGAAAGCTTGAAGAAATTTCTTGGATTCTGAGAGT 1875
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 1876 CTCGGTGAGACTGTTTTCACAGAAGGATGTGTGTGTACCCAAAGGCTTACACATCAAAAAGAA 1935
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLysLysIleLeuAspThr 320
Db 1936 ATGGATAAAATAAGTGAAGAAATTTAGAAGATTCACACTAGCCCTATCAAAATCTTGATACA 1995
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1996 GTTCATTCTTGTGAAGAGCAAGGAACTTCAAAAGATCACTGTGAACAACGTACAGGA 2055
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2056 AAAATGGAAACAAATGAAAAGAAAGTTTGTGTGTACTGAAAAGAAAGAACTGTGAGAAGCAAAA 2115
Qy 361 GluIleLysSerGlnLeuGluAenGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 2116 GAAATAAAATCACACTTACAGTGTAGAACCAAAAGTTAAATGGGAACAGAGCTCTGCGATGTG 2175
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleLeuSerTrpMetLysIleAlaCys 398
Db 2176 AGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 2229
RESULT 5
AX303171 2232 bp DNA linear PAT 30-NOV-2001
LOCUS AX303171
DEFINITION Sequence 491 from Patent WO0179286.
ACCESSION AX303171
VERSION AX303171.1 GI:17383661
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and Hepler,W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: WO 0179286-A 491 25-OCT-2001;
FEATURES
source CORIXA CORPORATION (US)
Location/Qualifiers
1..2232
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.66e-163 Length: 2232
Score: 2023.00 Matches: 393
Percent Similarity: 98.9% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 6 Gaps: 0
US-09-489-079-24 (1-398) x AX303171 (1-2232)
Qy 1 MetGlnLysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAla 20
Db 1036 ATGCAAAAGCTGTCCCAATAAAGCCTTGGAAATTTGAAAATGAACAAACATTTGAGAGCA 1095
Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTrpGluGluSerSerTrpAspSer 40
Db 1096 GATGAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCT 1155

Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1156 GAGAGTCTCTGTGAGACTGTTTTCACAGAAGATGTGTGTTTACCACAGGCTTCRCATCA 1215
Qy 61 LysGluIleAspLysIleAenGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1216 AAAGAAATAGATAAATAAATGGAATTTAGAAAGGCTCTCTGTTAAAGATGGTCTTCTG 1275
Qy 81 LysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1276 AAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGACATG 1335
Qy 101 GlnThrPheLysAlaGluProGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1336 CAAACTTTCAAAGCAGAGCCCTCCCGAGAAGCCATCTGCTTCGAGCCCTGCCATTGAAATG 1395
Qy 121 GlnLysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAlaAsp 140
Db 1396 CAAAGTCTGTTCCAAATTAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCAGAT 1455
Qy 141 GluIleLeuProSerGluSerLysGlnLysAspTrpGluGluSerSerTrpAspSerGlu 160
Db 1456 GAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAG 1515
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1516 AGTCTCTGTGAGACTGTTTTCACAGAAGATGTGTGTTTACCACAGGCTTCRCATCAAAA 1575
Qy 181 GluIleAspLysIleAenGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 1576 GAAATAGATAAATAAATGGAATTTAGAAAGCTCTCTGATAATGATGTTTCTGAG 1635
Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1636 GCTCCCTGCAGAAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGACATGCA 1695
Qy 221 ThrPheLysAlaGluProGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1696 ACTTTCAAAGCAGAGCCCTCCCGAGAAGCCATCTGCTTCGAGCCCTGCCATTGAAATGCA 1755
Qy 241 LysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAlaAspGln 260
Db 1756 AAGTCTGTTCCAAATTAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCAGATCAG 1815
Qy 261 MetPheProSerGluSerLysGlnLysAenValGluGluAenSerTrpAspSerGluSer 280
Db 1816 ATGTTCCCTTCAGAAATCAAAACAAAGAAAGTTTCTTGGGATTTCTGAGAGT 1875
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 1876 CTCGGTGAGACTGTTTTCACAGAAGGATGTGTGTGTACCCAAAGGCTTACACATCAAAAAGAA 1935
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLysLysIleLeuAspThr 320
Db 1936 ATGGATAAATAAAGTGAAGAAATTTAGAAGATTCACACTAGCCCTATCAAAATCTTGATACA 1995
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1996 GTTCATTCTTGTGAAGAGCAAGGAACTTCAAAAGATCACTGTGAACAACAGTACAGGA 2055
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2056 AAAATGGAAACAAATGAAAAGAAAGTTTGTGTGTACTGAAAAGAAAGAACTGTGAGAAGCAAAA 2115
Qy 361 GluIleLysSerGlnLeuGluAenGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 2116 GAAATAAATAAATGAGTGTAGAACCAAAAGTTAAATGGGAACAGAGCTCTGCGATGTG 2175
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleLeuSerTrpMetLysIleAlaCys 398
Db 2176 AGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 2229

AR283451 AR283451 3681 bp DNA linear PAT 10-APR-2003
LOCUS
DEFINITION Sequence 463 from patent US 6528054.
ACCESSION AR283451
VERSION AR283451.1 GI:29720278
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3681)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and Hepler,W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6528054-A 463 04-WAR-2003;
FEATURES Location/Qualifiers
source 1.3681
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.2e-162 Length: 3681
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 6 Gaps: 0
US-09-489-079-24 (1-398) x AR283451 (1-3681)
Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1308 ATGCAAAAGTCTGTCCCAATAAAGCCTTGAATTTGAAATAATGAACAAACATTGAGAGCA 1367
Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1368 GATGAGATATCTCCATCAGATCCAAACAAAGAGACTATGAGNAAGTCTTGGGATTCT 1427
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTTTACCCAGGCTRCRCATCAA 1487
Qy 61 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1488 AAAGAAATAGATAAATAAATGGAATAATGAAGGGTCTCCTGTAAAGATGGTCTTCTG 1547
Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGGCTAACTGCGGAATGAAGTTTCTATTCCAACTAAAGCCTTAGAATTGATGGACATG 1607
Qy 101 GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1608 CAAACTTTCAAAGCAGAGCCTCCCGAGAAGCCATCTGCCTCGAGCCTGCCATTGAAATG 1667
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1668 CAAAGTCTGTTCCTCAATAAAGCCTTGAATTTGAAGATGAACAAACATTGAGAGCAGAT 1727
Qy 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1728 GAGATATCTCCATCAGATCCAAACAAAGAGACTATGAGNAAGTCTTGGGATTCTGAG 1787
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTTTACCCAGGCTRCRCATCAAAA 1847
Qy 181 GluIleAspLysIleAsnGlyLysLeuGluGluSerProAsnAspGlyPheLeuLys 200
Db 1848 GAAATAGATAAATAAATGGAATAATGAAGAGTCTCCTGATAATGATGGTTTCTGAAG 1907
Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220

Db 1908 GCTCCCTGCAGAAATGAAAGTTTCTATTCCAACTAAAGCCCTTAGAATTGATGGACATGCAA 1967
Qy 221 ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1968 ACTTTCAAAGCAGAGCCTCCCGAGAAGCCATCTGCCTTCGAGCCTGCCATTGAAATGCAA 2027
Qy 241 LysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGln 260
Db 2028 AAGTCTGTTCCTCAATAAAGCCTTGAATTTGAAGATGAACAAACATTGAGAGCAGATCAG 2087
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 2088 ATGTTCCTTTCCAGATCAAAACAAAGAGAGTTGAAAGAAATTTCTGGGATTTCTGAGAGT 2147
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 2148 CTCCTGTGAGACTGTTTACAGAGAGGATGTGTGTATCCCAAGGCTACACATCAAAAGAA 2207
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 2208 ATGATATAAATAAGTGGAAATTTAGAAGATTTCACTAGCCTATCAAAATCTTGGATACA 2267
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 2268 GTTCATTCTTTGTGAAAGAGCAGCGAATCTTCAAAAAGATCACTGTGAACACGTCACAGGA 2327
Qy 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2328 AAAATGGACAAATGAAAGAAAGTGTGTGTACTGAAAGAAACCTGTGAGAGGCAAAA 2387
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 2388 GAAATATAATCAGACTTAGAGAACCAAAAGTTAAATGGGAAACAGAGCTCTGCAGTGTG 2447
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
Db 2448 AGGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 2501
RESULT 7
AR344219 AR344219 3681 bp DNA linear PAT 17-AUG-2003
LOCUS
DEFINITION Sequence 463 from patent US 6579973.
ACCESSION AR344219
VERSION AR344219.1 GI:33740119
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3681)
AUTHORS Yuqiu,J., Dillon,D.C., Mitcham,J.L., Xu,J. and Harlocker,S.L.
TITLE Compositions for the treatment and diagnosis of breast cancer and methods for their use
JOURNAL Patent: US 6579973-A 463 17-JUN-2003;
FEATURES Location/Qualifiers
source 1.3681
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.2e-162 Length: 3681
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 6 Gaps: 0
US-09-489-079-24 (1-398) x AR344219 (1-3681)
Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1308 ATGCAAAAGTCTGTCCCAATAAAGCCTTGAATTTGAAATAATGAACAAACATTGAGAGCA 1367

```
Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1368 GATGAGATACCTCCCATCAGAAATCAAAACAAAGGACATATGAAGAAAGTTCTTGGGATCT 1427

Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGTGAGACTGTTTCAAGAAAGGATGTGTGTTTACCACCAAGGCTRCRCATCAA 1487

Qy 61 LysGluLeuAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1488 AAGAAATAGATAAATAAATGAATAGAAAGGCTCTCTGTTAAAGATGGTCTTCTG 1547

Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGCTAACTCCGGAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTCGATGACATG 1607

Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1608 CAAACTTTCAAAGCAGAGCCCTCCGAGAAAGCCATCTGCCTTCGAGCCTGCCATTTGAAATG 1667

Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1668 CAAAGTCTGTTCNAATTAAGCCTTGGAATTAAGATGAACAAACATTCGAGAGCAGAT 1727

Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1728 GAGATACCTCCCATCAGAAATCAAAACAAAGGACATATGAAGAAAGTTCTTGGGATCTGAG 1787

Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGTGAGACTGTTTTCACAGAAAGGATGTGTGTTTACCACCAAGGCTRCRCATCAA 1847

Qy 181 GluLeuAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 1848 GAAATAGATAAATAAATAAATGAAGAAATTAAGAGCTCTCTGTAATATGATGGTCTTCTGAG 1907

Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1908 GCTCCCTGCAGAAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTCGATGACATGCAA 1967

Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1968 ACTTTCAAAGCAGAGCCCTCCGAGAAAGCCATCTGCCTTCGAGCCTGCCATTTGAAATGCAA 2027

Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
Db 2028 AAGTCTGTTCAAATTAAGCCTTGGAATTAAGAAATGAAGATGAACAAACATTCGAGAGCAGATCAG 2087

Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 2088 ATGTTCCCTTCAGATCAAAACAAAGAAAGTTGAAGAAATTCCTTGGGATTCGAGAGT 2147

Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 2148 CTCCTGTGAGACTGTTTTCACAGAAAGGATGTGTGTGTACCAAGGCTTACACATCAAAAGAA 2207

Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 2208 ATGGATAAATAAATGGAAGAAATTAAGATTCAACTAGCCTATCAAAATCTTGATATACA 2267

Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 2268 GTTCATTTCTTGGAAGAGCAGGGNACTTCAAAAGATCACTGTGAACACAGTACAGGA 2327

Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLeuLysLysLysLeuSerGluAlaLys 360
Db 2328 AAAATGGAACAAATGAAAGAAAGTTTGTGTACTGAAAAAGAAACTGTGCAAGACCAAAA 2387

Qy 361 GluIleLysSerGlnLeuGlnLysValLysTyrGluGlnGluLeuCysSerVal 380
Db 2388 GAAATTAATATCAGTTAGAGAACCAAAAGTTAAATGGGAACAAGAGCTTCGCAAGTGTG 2447

Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
```

```
Db 2448 AGGTTTCTCACTCATGAAATGAAATATATCTTTACATGAAAAATTGCATGT 2501

RESULT 8
LOCUS AR351420 3681 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 463 from patent US 6586572.
ACCESSION AR351420
VERSION AR351420.1 GI:33753099
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3681)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and Hepler,W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6586572-A 463 01-JUL-2003;
FEATURES Location/Qualifiers
source 1..3681
/mol_type="genomic DNA"

ALIGNMENT
Alignment Scores: 1.2e-162 Length: 3681
Pred. No.: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: Gaps: 0

US-09-489-079-24 (1-398) x AR351420 (1-3681)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1308 ATGCANAAGTCTGTCCCAATTAAGCCTTGGAAATGAATAATGAACAAACATTCGAGACA 1367

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1368 GATGAGATACCTCCCATCAGAAATCAAAACAAAGGACATATGAAGAAAGTTCTTGGGATCT 1427

Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGTGAGACTGTTTTCACAGAAAGGATGTGTGTTTACCACCAAGGCTRCRCATCAA 1487

Qy 61 LysGluLeuAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1488 AAGAAATAGATAAATAAATGAATGAAGGCTCTCTGTTAAAGATGGTCTTCTG 1547

Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGCTAACTCCGGAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTCGATGACATG 1607

Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1608 CAAACTTTCAAAGCAGAGCCCTCCGAGAAAGCCATCTGCCTTCGAGCCTGCCATTTGAAATG 1667

Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1668 CAAAGTCTGTTCNAATTAAGCCTTGGAATTAAGATGAACAAACATTCGAGAGCAGAT 1727

Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1728 GAGATACCTCCCATCAGAAATCAAAACAAAGGACATATGAAGAAAGTTCTTGGGATCTGAG 1787

Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGTGAGACTGTTTTCACAGAAAGGATGTGTGTTTACCACCAAGGCTRCRCATCAA 1847

Qy 181 GluLeuAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
```

```
Db      1848  GAAATAGATAAAATAAATGGAATAATTAGAAGAGTCTCTCTGATATGATGGTTTTCTCGAAG 1907
Qy      201  AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db      1908  GTCCTCCCTGCGAAGTAAAGTTTCTATTCCAACTAAAGCCTTTAGAAATTTGATGGACATCAA 1967
Qy      221  ThrPheLysAlaGluProCysGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db      1968  ACTTTCAAAGCAGAGCCTCCCGAAGGCCATCTGCCTTCGAGCTGCCATTTGAATGCAG 2027
Qy      241  LysSerValProLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGln 260
Db      2028  AAGTCTCTTCCAAATAAAGCCTTGGAAATGAAGAAATGAACAAACATTTGAGAGCAGATCAG 2087
Qy      261  MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db      2088  ATGTTCCTCTCAGAAATCAAAACAAAGAAAGTGAAGAAATTTCTGGGATTTCTGAGAGT 2147
Qy      281  LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db      2148  CTCGGTGAGACTGTTTCACAGAGGATGTGTGTACTCCCAAGGCTACACATCAAAAGAA 2207
Qy      301  MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db      2208  ATGGATAAAATAAGTGGAAATTTAGAAGATTTCAACTAGCCTATCAAAAATCTTGGATACA 2267
Qy      321  IleHisSerCysGluArgAlaArgGluLysAspHisCysGluGlnCysThrGly 340
Db      2268  GTTCATTTCTTGTGAAGACGACGGAAGTCTCAAAAAGATCACTGTGAACACGTCACGGA 2327
Qy      341  LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db      2328  AANAATGGAACAAATGAAGAAGTTTGTGTACTGAAAAAGAACTGTGAGAGCAAAA 2387
Qy      361  GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db      2388  GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2447
Qy      381  ArgPheLeuThrLeuMetLysMetLysIleIleSerTrpMetLysIleAlaCys 398
Db      2448  AGGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 2501

RESULT 9
AR454000 3681 bp DNA linear PAT 20-FEB-2004
LOCUS
DEFINITION Sequence 463 from patent US 6680197.
ACCESSION AR454000
VERSION AR454000.1 GI:42686790
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 3681)
AUTHORS Jiang, Y., Dillon, D. C., Mitcham, J. L., Xu, J., Harlocker, S. L.,
Hepler, W. T. and Henderson, R. A.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6680197-A 463 20-JAN-2004;
FEATURES Location/Qualifiers
source 1. .3681
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.2e-162 Length: 3681
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-24 (1-398) x AR454000 (1-3681)
```

```
Qy      1  MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db      1308  ATGCAAAGTCTGTGCCAAATAAAGCCTTGGAATTTGAAAATGTAACAAACATTTGAGACGA 1367
Qy      21  AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db      1368  GATGAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAGTCTCTTGGGATTTCT 1427
Qy      41  GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db      1428  GAGAGTCTCTGTGAGACTGTTTCACAGAGGATGTGTGTTTACCACAAAGGCTCRCAATCAA 1487
Qy      61  LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db      1488  AAGAAATAGATTAATTAATGGAATTAAGAGGGTCTCTGTTAAAGATGGTCTTCTG 1547
Qy      81  LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db      1548  AAGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATG 1607
Qy      101  GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db      1608  CAAACTTTCAAAGCAGAGCCTCCCGAAGCCATCTGCCTTCGAGCCTGCCATTTGAAATG 1667
Qy      121  GlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAsp 140
Db      1668  CAAAGTCTGTTCCAAATAAAGCCTTGGAAATTAAGAAATGAACAAACATTTGAGAGCAGAT 1727
Qy      141  GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db      1728  GAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAGTCTCTTGGGATTTCTGAG 1787
Qy      161  SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db      1788  AGTCTCTGTGAGACTGTTTCACAGAGGATGTGTGTTTACCACAAAGGCTCRCAATCAA 1847
Qy      181  GluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db      1848  GAAATAGATAAAATTAATGGAATTAAGAAATTAAGAGAGTCTCTCTGATATGATGGTTTCTGAG 1907
Qy      201  AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db      1908  GTCCTCCCTGCGAAGTAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATGCA 1967
Qy      221  ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db      1968  ACTTTCAAAGCAGAGCCTCCCGAAGGCCATCTGCCTTCGAGCCTGCCATTTGAAATGCA 2027
Qy      241  LysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGln 260
Db      2028  AAGTCTGTTCCAAATAAAGCCTTGGAAATTAAGAAATGAACAAACATTTGAGAGCAGATCAG 2087
Qy      261  MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db      2088  ATGTTCCTCTCAGAAATCAAAACAAAGAAAGTGAAGAAATTTCTTGGGATTTCTGAGAGT 2147
Qy      281  LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db      2148  CTCGGTGAGACTGTTTCACAGAGGATGTGTGTACTCCCAAGGCTACACATCAAAAGAA 2207
Qy      301  MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db      2208  ATGGATAAAATAAGTGGAAATTTAGAAGATTTCAACTAGCCTATCAAAAATCTTGGATACA 2267
Qy      321  IleHisSerCysGluArgAlaArgGluLysAspHisCysGluGlnCysThrGly 340
Db      2268  GTTCATTTCTTGTGAAGACGACGGAAGTCTCAAAAAGATCACTGTGAACAAACGTCACGGA 2327
Qy      341  LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db      2328  AAAATGGAAACAAATGAAAGAAAGTTTTGTGTACTGAAAGAAAGAACTGTGAGAGCAAAA 2387
```

Db	1788	AGTCTCTGTGAGACTGTTTTCACAGAAGGATGTGTGTTTACCCCAAGGCTCRCATCAAAA	184
Qy	181	GluileAspIysIleAsnGlyIysLeuGluGluSerProAspAenAepGlyPheLeuLys	200
Db	1848	GAATAATAGATAAAATAAATGCAAAATTTAGAAAGAGTCTCTGATAATGATGGTTTCTCAAG	1907
Qy	201	AlaProCysArgMetIysValSerIleProThrIysAlaLeuGluLeuMetAspMetGln	220
Db	1908	GCTCCCTCGCAGAAATGAAGTTCTTATTCACAACTAAAGCCCTTAGAATTTGATGGACATGC	1967
Qy	221	ThrPheIysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln	240
Db	1968	ACTTTCAAGCAGAGCTCCCGAAGACCATCTCGCTTCAGGCTGCCATTGAATGC	2027
Qy	241	LysSerValProAenIysAlaLeuGluLeuLysAsnGluGlnThrLeuA-gAlaAspGln	260
Db	2028	AAGTCTGTTCCAAATAAAGCCTTTGGAATTCGAAGATGAACAAACATTTGAGAGCAGATCAG	2087
Qy	261	MetPheProSerGluSerIysGlnLysAsnValGluGluAenSerTrpAenSerGluSer	280
Db	2088	ATGTTCCCTTTCAGAAATCAAAACAAAGAGAGGTTGAAGAAATTTCTGGGATTTCTGAGAGT	2147
Qy	281	LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu	300
Db	2148	CTCGTGAGACTGTTTTCACAGAAGGATGTGTGTGTACCCAAAGGCTACACATCAAAAAGAA	2207
Qy	301	MetAspIysIleSerGlyIysLeuGluAspSerThrSerLeuSerIysIleLeuAspThr	320
Db	2208	ATGGATAAAATAAGTCGNAATTTAGAAAGATTCACATAGCCTATCAAAAATCTTCGATACA	2267
Qy	321	IleHisSerCysGluArgAlaArgGluLeuGlnIysAspHisCysGluGlnCysThrGly	340
Db	2268	GTTTCATTTCTTGTAAGAGAGCAGCGGAACATTCAAAAGAGATCACCTGTGTGAACAACTGACAGGA	2327
Qy	341	LysMetGluGlnMetIysIysLysPheCysValLeuLysIysLysLeuSerGluAlaLys	360
Db	2328	AAATGGAACAAATGAATAAAGAGTTTGTGTACTGAAAAGAACTGTGAGAAGCAAA	2387
Qy	361	GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal	380
Db	2388	GAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG	2447
Qy	381	ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetIysIleAlaCys	398
Db	2448	AGGTTTCTCACACTCATGAAATGAAATATATCTTTACATGAAATTTGCATGT	2501
RESULT 11			
LOCUS	AX282970	3681 bp	DNA
DEFINITION	Sequence 19 from Patent WO0175171.		
ACCESSION	AX282970		
VERSION	AX282970.1	GI:16609903	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Houghton,R.L., Dillon,D.C., Molesh,D.A., Xu,J., Zehentner,B. and Persing,D.H.		
TITLE	Methods, compositions and kits for the detection and monitoring of breast cancer		
JOURNAL	Patent: WO 0175171-A 19 11-OCT-2001; CORIXA CORPORATION (US)		
FEATURES	Location/Qualifiers		
source	1. .3681		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:	1.2e-162	Length:	3681
Pred. No.:			

Score:	2023.00	Matches:	393	
Percent Similarity:	98.99%	Conservative:	1	
Best Local Similarity:	98.74%	Mismatches:	4	
Query Match:	98.59%	Indels:	0	
DB:	6	Gaps:	0	
US-09-489-079-24 (1-398) x AX282970 (1-3681)				
Qy	1	MetGlnLysSerValProAsnLysAlaLeuLysAsnGluThrLeuArgAla	20	
Db	1308	ATGCANAAGTCTGTCCCAATAAAGCCTTGGAATTGAAAAATGAACAAACATTGAGAGCA	1367	
Qy	21	AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer	40	
Db	1368	GATGAGATCTCCCATCAGATCCAAACAAAGAGACTATGAGAAAGTTCTTGGGATTCT	1427	
Qy	41	GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln	60	
Db	1428	GAGAGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTACCAAGGCTRCRCATCAA	1487	
Qy	61	LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu	80	
Db	1488	AAAGAAATAGATAAAATAAATGGAATAATGAAGGGTCTCCTGTTAAAGATGGTCTTCTG	1547	
Qy	81	LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet	100	
Db	1548	AAGCTTAACCTGCGGAATGAAAGTTCTATTCCAATTAAGCCTTAGAATTGATGACATG	1607	
Qy	101	GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet	120	
Db	1608	CAAACTTTCAAAGCAGAGCCTCCGAGAGCCATCTGCTGATAATGATGGTTTCTGAAG	1667	
Qy	121	GlnLysSerValProAsnLysAlaLeuLysAsnGluThrLeuArgAlaAsp	140	
Db	1668	CAAAAGTCTGTTCCAATAAAGCCTTCGAAATGAAGAAATGAACAAACATTGAGAGCAGAT	1727	
Qy	141	GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu	160	
Db	1728	GAGATATCTCCATCAGAAATCCAAACAAAGAGACTATGAGAAAGTTCTTGGGATTCTGAG	1787	
Qy	161	SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys	180	
Db	1788	AGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTATTACCAAGGCTRCRCATCAAAA	1847	
Qy	181	GluIleAspLysIleAsnGlyLysLeuGluGlySerProAspAsnAspGlyPheLeuLys	200	
Db	1848	GAATATAGATAAAATAAATGGAATAATGAAGAGTCTCCTGATAATGATGGTTTCTGAAG	1907	
Qy	201	AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln	220	
Db	1908	GTCCTCTGCAGATCAAAAGTTTCTATTCCAATTAAGCCTTAGAATTGATGGACATGCAA	1967	
Qy	221	ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln	240	
Db	1968	ACTTTCAAAGCAGAGCCTCCGAGAGCCATCTGCTTCGAGCCTGCCATTGAAATGCAA	2027	
Qy	241	LysSerValProAsnLysAlaLeuLysAsnGluThrLeuArgAlaAspGln	260	
Db	2028	AAGTCTCTTCCAAATAAAGCCTTCGAAATGAAGAAATGAACAAACATTGAGAGCAGATCAG	2087	
Qy	261	MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer	280	
Db	2088	ATGTTCTCTTCAGANTCAAACAAAGAGAGTTGAGAGAAATTTCTTGGGATTCTGAGAGT	2147	
Qy	281	LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu	300	
Db	2148	CTCCGTGAGACTGTTTACAGAGAGATGTGTGTACCCCAAGGCTACACATCAAAAAGAA	2207	
Qy	301	MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr	320	
Db	2208	ATGATATAAATAAAGTGAATAATGAAGATTCAACTAGCCTATCAAAAATCTTGGATACA	2267	
Qy	321	IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly	340	
Db				
Db	2268	GTTCATCTTGTGAAAGAGCAGGGAACCTTCAAAAGATCACTGTGAACACGTACAGGA	2327	
Qy	341	LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys	360	
Db	2328	AAAATGGAACAAATGAAAAAGAGTTTGTGTACTGAAAAAGAACTGTTCAGAAAGCAAAA	2387	
Qy	361	GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnLeuCysSerVal	380	
Db	2388	GAAATAAATACACAGTTAGAGAACCAAAAAGTTTAAATGGGAACAAAGACTCTCGAGTGTG	2447	
Qy	381	ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys	398	
Db	2448	AGGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAAAATTGCATGT	2501	
RESULT 12				
AX303143 3681 bp DNA linear PAT 30-NOV-2001				
LOCUS				
DEFINITION Sequence 463 from Patent WO0179286.				
ACCESSION AX303143				
VERSION AX303143.1 GI:17383644				
KEYWORDS				
SOURCE Homo sapiens (human)				
ORGANISM				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE				
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and				
Hepler, W.T.				
TITLE Compositions and methods for the therapy and diagnosis of breast				
cancer				
JOURNAL Patent: WO 0179286-A 463 25-OCT-2001;				
CORIXA CORPORATION (US)				
FEATURES				
source 1. 3681				
Location/Qualifiers				
/organism="Homo sapiens"				
/mol_type="unassigned DNA"				
/db_xref="taxon:9606"				
ORIGIN				
Alignment Scores:				
Pred. No.:	1.2e-162	Length:	3681	
Score:	2023.00	Matches:	393	
Percent Similarity:	98.99%	Conservative:	1	
Best Local Similarity:	98.74%	Mismatches:	4	
Query Match:	98.59%	Indels:	0	
DB:	6	Gaps:	0	
US-09-489-079-24 (1-398) x AX303143 (1-3681)				
Qy	1	MetGlnLysSerValProAsnLysAlaLeuLysAsnGluThrLeuArgAla	20	
Db	1308	ATGCANAAGTCTGTCCCAATAAAGCCTTGGAATTGAAAAATGAACAAACATTGAGAGCA	1367	
Qy	21	AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer	40	
Db	1368	GATGAGATATCTCCATCAGAAATCCAAACAAAGAGACTATGAGAAAGTTCTTGGGATTCT	1427	
Qy	41	GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln	60	
Db	1428	GAGAGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTACCAAGGCTRCRCATCAA	1487	
Qy	61	LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu	80	
Db	1488	AAAGAAATAGATAAAATAAATGGAATAATGAAGGGTCTCCTGTTAAAGATGGTCTTCTG	1547	
Qy	81	LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet	100	
Db	1548	AAGCTTAACCTGCGGAATGAAAGTTCTATTCCAATTAAGCCTTAGAATTGATGACATG	1607	
Qy	101	GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet	120	
Db	1608	CAAACTTTCAAAGCAGAGCCTCCGAGAGCCATCTGCTGATAATGATGGTTTCTGAATG	1667	

Qy	121	GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp	140
Db	1668	CAAAAGTCTGTTCCAAATAAAGCCCTTGGAAATGAAGATGAACAAACATTGAGAGCAGAT	1727
Qy	141	GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu	160
Db	1728	GAGATACCTCCCATCAGAAATCAAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAG	1787
Qy	161	SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys	180
Db	1788	AGTCTCTGTGAGACTGTTTCACAGAGGATGTGTGTTTACCAAGGCTTRCRAATCAAAA	1847
Qy	181	GluLeuAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys	200
Db	1848	GAATAGATAAATAAATGAAGAAATTAGAAGACTCTCTGATAATGATGGTTTCTGAAG	1907
Qy	201	AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln	220
Db	1908	GCTCCCTGCAGAAATGAAGTTTCTATCCAACTAAAGCCTTAGAAATGATGACATGCCA	1967
Qy	221	ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln	240
Db	1968	ACTTTCAAAGCAGAGCCCTCCGAGAGCCATCTGCCTTCGAGCCCTGCCATTTGAATGCCA	2027
Qy	241	LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln	260
Db	2028	AAGTCTGTTCNAATAAAGCCCTTGGAAATGAAGATGAACAAACATTGAGAGCAGATCAG	2087
Qy	261	MetPheProSerGluSerLysGlnLysAsnValGluGluAenSerTrpAspSerGluSer	280
Db	2088	ATGTTCCCTTCAGAAATCAAAACAAAGAAAGTTGAAGAAATTTCTTGGGATTTCTGAGAGT	2147
Qy	281	LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu	300
Db	2148	CTCCGTGAGACTGTTTTCACAGAGGATGTGTGTACCCCAAGGCTACACATCAAAAGAA	2207
Qy	301	MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr	320
Db	2208	ATGGATAAATAAATGGAATAATAGAGATTCACACTAGCCTATCAAAATCTTGGATACA	2267
Qy	321	IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly	340
Db	2268	GTTTCATTTCTGTGAAGAGCAGAGGAACTTCAAAAGATCACTGTGAACAAACGTACAGGA	2327
Qy	341	LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys	360
Db	2328	AAAATGGAAACAATGAAAAGAAAGTTTGTGTACTGAAAAGAAACTGTGAGAACGAAAA	2387
Qy	361	GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal	380
Db	2388	GAATATAATACAGTTAGAGAACCCAAAGTTAATGGGAACAGAGCTCTGCAGTGTG	2447
Qy	381	ArgPheLeuThrLeuMetLysMetLysIleLeuSerTyrMetLysIleAlaCys	398
Db	2448	AGTTTCTCACACTCATGAAATGAAATTAATCTCTTACATGAAAATTCATGT	2501
RESULT 13			
AR283467			
LOCUS			
Sequence 490 from patent US 6528054.			
AR283467			
AR283467.1 GI:29720294			
KEYWORDS			
SOURCE			
ORGANISM			
Unknown.			
Unclassified.			
1 (bases 1 to 3288)			
Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and			
Hepler, W.T.			
Compositions and methods for the therapy and diagnosis of breast			
cancer			
Patent: US 6528054-A 490 04-MAR-2003;			

FEATURES		Location/Qualifiers	
source		1..3288	
		/organism="unknown"	
		/mol_type="genomic DNA"	
ORIGIN			
Alignment Scores:			
Pred. No.:		8.38e-158	Length
Score:		1965.50	Matches
Percent Similarity:		85.65%	Cons
Best Local Similarity:		85.43%	Mismat
Query Match:		95.78%	Indel
DB:		6	Gaps:
US-09-489-079-24 (1-398) x AR283467 (1-3288)			
Qy	1	MetGlnLysSerValProAsnLysAlaLeu	LysAlaLeu
Db	1036	ATGCCAAAGTCTGTCCTCCCAATAAAGCCTTG	
Qy	21	AspGluIleLeuProSerGluSerLysGln	LysGln
Db	1096	GATGAGATACTCCCATCAGAAATCCAAACAA	
Qy	41	GluSerLeuCysGluThrValSerGlnLys	LysGln
Db	1156	GAGAGTCTCTGTGAGACTGTTTTTCACAGA	
Qy	61	LysGluIleAspLysIleAsnGlyLysLeu	LysLeu
Db	1216	AAAGAAATAGATAAATAAATGGAATAATT	
Qy	81	LysAlaAsnCysGlyMetLysValSerIle	LysIle
Db	1276	AAGCTTAAGTCTCGGAATGAAAGTTTCTAT	
Qy	101	GlnThrPheLysAlaGluProGluLys	LysGlu
Db	1336	CAAACTTTCAAAGCAGAGCGCTCCCGAGA	
Qy	121	GlnLysSerValProAsnLysAlaLeuGlu	LysGlu
Db	1396	CAAAAGTCTGTTTCCAAATAAAGCCTTGG	
Qy	141	GluIleLeuProSerGluSerLysGlnLys	LysGln
Db	1456	GAGATACTCCCATCAGAAATCCAAACAA	
Qy	161	SerLeuCysGluThrValSerGlnLysAsp	LysAsp
Db	1516	AGTCTCTGTGAGACTGTTTTTCACAGA	
Qy	181	GluIleAspLysIleAsnGlyLysLeuGlu	LysLeu
Db	1576	GAATAGATAAATAAATGGAATAATTAG	
Qy	201	AlaProCysArgMetLysValSerIlePro	LysIlePro
Db	1636	GCTCCCTGCAGAAATGAAAGTTTCTATT	
Qy	221	ThrPheLysAlaGluProGluLysPro	LysPro
Db	1696	ACTTTCAAAGCAGAGCGCTCCCGAGA	
Qy	241	LysSerValProAsnLysAlaLeuGluLeu	LysLeu
Db	1756	AAGTCTGTTTCCAAATAAAGCCTTTGG	
Qy	261	MetPheProSerGluSerLysGlnLysAsn	LysAsn
Db	1816	ATGTTCCCTTCAGAAATCAAAACAA	
Qy	281	LeuArgGluThrValSerGlnLysAspVal	LysAspVal

```
Db 1876 CTCGGTGAGACTGTTTACAGAGGATGTGTGTGTACCCAGGCTACACATCAAAAGAA 1935
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 1936 ATGGATAAAATAAGTGAATAATTAGAGATTCAACTAGCCCTATCAAAAATCTTGGATACA 1995
Qy 321 IleHisSerCysGluAArgAlaAArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1996 GTTCATTCTTGTGAAAGAGCAGGGAAGCTTCAAAAAGATCACTGTGAACAACGTACAGGA 2055
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLysLeuSerGluAlaLys 360
Db 2056 AAAATGCAACAAATGAAAAGAGTGTGTGTACTGAAAAAGAACTGTGAGAAGCAAAA 2115
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysLysValLysTrpGluGlnGluCysSerVal 380
Db 2116 GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2175
Qy 380 ----- 380
Db 2176 AGATTGACTTTAAACCAAGAAGAAGAGAGAGAAATGCCGATATATTAAATGAAAAA 2235
Qy 380 ----- 380
Db 2236 ATTAGGAAGAAATTAGGAAGAATCGAAGAGCAGCATAGGAAGAGTTAGAAAGTGAACAA 2295
Qy 380 ----- 380
Db 2296 CAACCTGAACAGCGCTCTCAGAATAACAAGATATAGAATTGAAGAGTGTGAGAAAGTAATTG 2355
Qy 381 ----ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
Db 2356 AATCAGGTTTCTCACTCATGAAATGAAATATATCTCTTACATGAAATTCATGT 2413

RESULT 14
AR454016
LOCUS AR454016 3288 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 490 from patent US 6680197.
ACCESSION AR454016
VERSION AR454016.1 GI:42686806
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3288)
AUTHORS Jiang, Y., Dillon, D. C., Mitcham, J. L., Xu, J., Harlocker, S. L.,
Hepler, W. T. and Henderson, R. A.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6680197-A 490 20-JAN-2004;
FEATURES Location/Qualifiers
source 1..3288
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 8.38e-158 Length: 3288
Score: 1965.50 Matches: 393
Percent Similarity: 85.65% Conservative: 1
Best Local Similarity: 85.43% Mismatches: 4
Query Match: 95.78% Indels: 62
DB: 6 Gaps: 1

US-09-489-079-24 (1-398) x AR454016 (1-3288)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1036 ATGCCAAAGTCTGTCCCAATTAAGCCCTTGGAAATGAAATGAAACAACATTTGAGAGCA 1095
Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
```

```
Db 1096 GATGAGATACCTCCCATCAGAATCAAAACAAAAGGACTATGAAGAAAGTTCTTGGGATTCT 1155
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProCysAlaAlaHisGln 60
Db 1156 GAGAGTCTCTGTGAGACTGTTTACAGAGGATGTGTGTGTACCCAGGCTCRCATCAA 1215
Qy 61 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1216 AAGAANAATAGTAANAATAAATGGAANAATTAGAAGGGTCTCTCTGTTAAAGATGGTCTCTG 1275
Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1276 AAGCTAACTGCGGAATGAAAGTTTCTATTCCAACATAAGCCTTAGAATTTGATGACATG 1335
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1336 CAAACTTTCAAAGCAGAGCCCTCCGAGAACCCATCTGCCCTTCGAGCGTCCATTGAAATG 1395
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1396 CAAAGTCTGTTCCAAATAAAGCTTGGAAATTGAAGATGAACAACATTTGAGAGCAGAT 1455
Qy 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1456 GAGATACCTCCCATCAGAATCAAAACAAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAG 1515
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1516 AGTCTCTGTGAGACTGTTTACAGAGAGGATGTGTGTATCCCAAGGCTCRCATCAAAA 1575
Qy 181 GluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 1576 GAAATAGATAAANAATAATGGAANAATTAGAAGAGTCTCTGTAATGATGTTCTGAG 1635
Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1636 GCTCCCTGCGAGAATGAAAGTTTCTATTCCAACATAAGCCTTAGAATTTGATGAGCATGCA 1695
Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1696 ACTTTCAAAGCAGAGCGCTCCCGAGAGGCCATCTGCCCTTCGAGCGTCCATTGAAATGCA 1755
Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
Db 1756 AAGTCTGTTCCAAATAAAGCTTGGAAATTGAAGATGAACAACATTTGAGAGCAGATCAG 1815
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 1816 ATGTTCCCTTCAGAAATCAAAACAAAAGAGSGTTGAAGAAAATTTCTTGGGATTTCTGAGAGT 1875
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 1876 CTCCTGTGAGACTGTTTACAGAGAGGATGTGTGTATCCCAAGGCTACACATCAAAAAGAA 1935
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 1936 ATGGATAAANAATAGTGGANAATTAGAAGATTCACCTAGCCCTATCAAAAATCTTGGATACA 1995
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1996 GTTCATTCTTGTGAAAGAGCAGGGAACCTTCAAAAAGATCACTGTGAACAACGTCACAGGA 2055
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2056 AAAATGCAACAAATGAAAAGAGTGTGTGTACTGAAAAAGAACTGTGAGAAGCAAAA 2115
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluCysSerVal 380
Db 2116 GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2175
Qy 380 ----- 380
Db 2176 AGATTGACTTTAAACCAAGAAGAAGAGAGAAATGCCGATATATTAAATGAAAAA 2235
```

QY 380 ----- 380

Db 2236 ATTAGGAAGATTAGGAAGATCGAAGCAGCATAGGAAGAGTTAGAAAGTGAACAA 2295

QY 380 ----- 380

Db 2296 CAACCTTGACAGGCTCTCAGATACAGATATAGAAATTTGAAGAGTGTTAGAAAGTAAATTTG 2355

QY 381 ----ArgPheLeuThrLeuMetLysMetLysAlaIleSerTyrMetLysIleAlaCys 398

Db 2356 AATCAGGTTTTCACACTCATGAAATGAAATTAATCTCTTACATGAAATTCATGT 2413

RESULT 15

AX303170 3288 bp DNA linear PAT 30-NOV-2001

LOCUS AX303170

DEFINITION Sequence 490 from Patent WO0179286.

ACCESSION AX303170

VERSION AX303170.1 GI:17383660

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and Hepler, W.T.

TITLE Compositions and methods for the therapy and diagnosis of breast cancer

JOURNAL Patent: WO 0179286-A 490 25-OCT-2001;

FEATURES

source 1. 3288

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 8,388-158 Length: 3288

Score: 1965.50 Matches: 393

Percent Similarity: 85.65% Conservative: 1

Best Local Similarity: 85.43% Mismatches: 4

Query Match: 95.78% Indels: 62

DB: 6 Gaps: 1

US-09-489-079-24 (1-398) x AX303170 (1-3288)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20

Db 1036 ATGCATAAGTCTGTCCCAATTAAGCCTTGGAAATTTGAAATTTGAATAATGAACAAACATTGAGAGCA 1095

QY 21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40

Db 1096 GATGAGATACTCCCATCAGATCCAAACAAAGAGGACTATGAAGAAAGTTCTTTGGGATTCT 1155

QY 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60

Db 1156 GAGAGTCTCTGTGAGACTGTTTCACAGAAGGAGTGTGTTTACCCAGGCTRCRCATCAA 1215

QY 61 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80

Db 1216 AAGAATAATAGATAATAATGGAATAATTAGAAGGCTCTCCTGTTAAAGATGTTCTTCTG 1275

QY 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100

Db 1276 AAGGCTAACTGGCGGATGAAGTTTCTATTCAACTAAAGCCTTAGAAATTTGATGACATG 1335

QY 101 GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120

Db 1336 CAACATTTCAAAGCAGAGGCTCCCGAGAAGGCATCTGCCTTCGAGCCTCCCATTTGAATG 1395

QY 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140

Search completed: August 2, 2005, 02:51:01
Job time : 6061.64 secs

Db 1396 CAAAGTCTGTTCATAATAAGCCTTGGAAATTTGAAGATTAAGAAATGAACAAACATTGAGAGCAT 1455

QY 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160

Db 1456 GAGATATCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTCTCGAG 1515

QY 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180

Db 1516 AGTCTCTGTGAGACTGTTTTCACAGAAGGATGTGTGTTTACCAAGGCTRCRCATCAAAA 1575

QY 181 GluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200

Db 1576 GAAATAGATAATAATAATGGAATAATTAGAAGAGTCTCTGTATAATGATGGTTTTCTGAAG 1635

QY 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220

Db 1636 GCTCCTGCAGATGAAGTTTCTATCCNACTAAAGCCTTAGAATTTGATGACATGCAA 1695

QY 221 ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln 240

Db 1696 ACTTTCAAAGCAGAGCCTCCCGAGAAGCATCTGCCTTCGAGCCTGCCATTTGAAATGCAA 1755

QY 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260

Db 1756 AAGTCTGTTCATAATAAGCCTTGGAAATTTGAAGATTAAGAAATGAACAAACATTGAGAGCATCAG 1815

QY 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280

Db 1816 ATGTTCCCTTCAGATCCAAACAAAGAGSCTTGAAGAAATTTCTTTGGGATTCTCGAGT 1875

QY 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300

Db 1876 CTCGTGAGACTGTTTTCACAGAAGGATGTGTGTGTACCAAGGCTACACATCAAAAAGAA 1935

QY 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320

Db 1936 ATGGATAATAATAAGTGAATAATTAGAAGATTCAACTAGCCTATCAAAAATCTTGATACA 1995

QY 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340

Db 1996 GTTCATCTTCTGTGAAGAGCAAGGGAACCTTCAAAAGATCACTGTGAACACAGTACAGA 2055

QY 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360

Db 2056 AAAATGGAACAAATGAAAAAGAGTTTGTGTACTGAAAAAGAAACTGTTCAGAGCAAAA 2115

QY 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTyrGluGlnGluLeuCysSerVal 380

Db 2116 GAAATAATAATCACAGTTAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2175

QY 380 ----- 380

Db 2176 AGATTGACTTTAAACCAAGAGAAGAGAAGAAATGCCGATATATTAAATGAACAA 2235

QY 380 ----- 380

Db 2236 ATTAGGAAGAAATTAGGAAGAAATCGAAGCAGCATAGGAAGAGTTAGAAAGTGAACAA 2295

QY 380 ----- 380

Db 2296 CAACCTTGACAGGCTCTCAGATACAGATATAGAAATTTGAAGAGTGTTAGAAAGTAAATTTG 2355

QY 381 ----ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398

Db 2356 AATCAGGTTTTCACACTCATGAAATGAAATTAATCTCTTACATGAAATTCATGT 2413

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:23:28 ; Search time 737.055 Seconds
(without alignments)
3196.585 Million cell updates/sec

Title: US-09-489-079-24

Perfect score: 2052

Sequence: 1 MQKSPVKALBKNEQTLRA.....SVRFLLMKMKIISYMKIAC 398

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xl
-Q=/cgn2_1/USPTO.spool_p/US09489079/runat_01082005_141249_13290/app.query.fasta_1.1635
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN_1_1418 @runat_01082005_141249_13290 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2052	100.0	2683	AA59015	AA59015 Nucleotid
2	2052	100.0	2683	AA59014	AA59014 Nucleotid
3	2034	98.6	3681	AA547405	AA547405 Human cDN
4	2023	98.6	2232	AA547422	AA547422 Human cDN
5	2023	98.6	2232	AB564023	AB564023 Human bre

6	2023	98.6	2232	10	ABT33235	Human tum
7	2023	98.6	2232	11	ADL93153	Human bre
8	2023	98.6	3681	4	AAFL17975	Human bre
9	2023	98.6	3681	4	AAI67218	B726P spl
10	2023	98.6	3681	6	ABS64006	Human bre
11	2023	98.6	3681	10	ABT33218	Human tum
12	2023	98.6	3681	11	ADL93125	Human bre
13	2023	98.6	3681	12	ADE44415	Human cDN
14	2011	98.0	1953	10	ABT33259	Human tum
15	2011	98.0	1953	11	ADL93211	Human bre
16	1965.5	95.8	3288	4	AA547421	Human cDN
17	1965.5	95.8	3288	6	ABS64022	Human bre
18	1965.5	95.8	3288	10	ABT33234	Human tum
19	1965.5	95.8	3288	11	ADL93152	Human bre
20	1965.5	95.8	3865	4	AA547411	Human cDN
21	1965.5	95.8	3865	6	ABS64012	Human bre
22	1965.5	95.8	3865	10	ABT33224	Human tum
23	1965.5	95.8	3865	11	ADL93136	Human bre
24	1965.5	95.8	3865	12	ADE44426	Human cDN
25	1963.5	95.7	4047	10	ABT33272	Human tum
26	1963.5	95.7	4047	11	ADL93228	Human bre
27	1963.5	95.7	4458	6	ABT07694	Breast ca
28	1963.5	95.7	4458	8	ACC50246	Breast ca
29	1963.5	95.7	4458	10	ABT33271	Human tum
30	1963.5	95.7	4458	11	ADL93226	Human bre
31	1960.5	95.5	3045	10	ABT33258	Human tum
32	1960.5	95.5	3045	11	ADL93210	Human bre
33	1948	94.9	2307	4	AAFL17980	Human spl
34	1948	94.9	2307	4	AAI67223	B726P spl
35	1948	94.9	2307	4	AA547410	Human cDN
36	1948	94.9	2307	6	ABS64011	Human bre
37	1948	94.9	2307	10	ABT33223	Human tum
38	1948	94.9	2307	11	ADL93130	Human bre
39	1948	94.9	2307	12	ADE44420	Human cDN
40	1661.5	81.0	1337	4	AAFL17979	Human bre
41	1661.5	81.0	1337	4	AAI67222	B726P spl
42	1661.5	81.0	1337	4	AA547409	Human cDN
43	1661.5	81.0	1337	6	ABS64010	Human bre
44	1661.5	81.0	1337	10	ABT33222	Human tum
45	1661.5	81.0	1337	11	ADL93129	Human bre

ALIGNMENTS

RESULT 1
AAAS9015
ID AAAS9015 standard; cDNA; 2683 BP.
AC AAAS9015;
DT 07-NOV-2000 (first entry)
XX Nucleotide sequence of BS322 consensus cDNA sequence.
DE BS322; breast tissue marker; breast disease; breast cancer; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 57..1253
FT /*tag= a
FT /note= "encodes AAB07638"
FT CDS 1172..2125
FT /*tag= b
FT /note= "encodes AAB07639"
XX WO200043420-A1.
XX 27-JUL-2000.
XX 21-JAN-2000; 2000WO-US001452.
XX 21-JAN-1999; 99US-00234716.

XX PA (ABBO) ABBOTT LAB.

XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
XX PI Stroupe SD;

XX DR WPI: 2000-499217/44.

XX DR P-PSDB; AAB07638, AAB07639.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
PT detection and diagnosis of breast disease especially breast cancer.

XX PT Claim 12; Page 119-120; 126pp; English.

XX CC The present sequence encodes a human BS322 polypeptide. BS322 is a breast
CC tissue marker. The BS322 polynucleotides and polypeptides are used to
CC detect and diagnose breast disease, e.g. breast cancer. The BS322
CC polynucleotides are useful as a source of probes and primers, and the
CC BS322 polypeptides are useful as antigens

XX SQ Sequence 2683 BP; 1043 A; 462 C; 502 G; 676 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.94e-176 Length: 2683
Score: 2052.00 Matches: 398
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-489-079-24 (1-398) x AAAS9015 (1-2683)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlnThrLeuArgAla 20
DB 57 ATGCAGAAAGTCTGTTCACAAATAAGAGCTTGGAAATGAAGAATGAACAAACATTGAGAGCA 116

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
DB 117 GATGAGATACTCCCATCAGATCCAAACAAAGAGACTATGAGAAAGTTCTTGGGATTCT 176

QY 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
DB 177 GAGAGTCTCTGTGAGAGCTGTTTACAGAGGATGTGTGTACCAAGGCTGCAGCATCAA 236

QY 61 LysGluLeuAspLysIleAsnGlyLysLeuGluClySerProValLysAspGlyLeuLeu 80
DB 237 AAAGAAAATAGATAAAATAAATGGAAAATTAAGAGGGTCTCTGTAAAGATGGTCTTCTG 296

QY 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
DB 297 AAGCTTAACCTCGGAATGAAGTTCTTATCCAACTAAGCCTTAGAATTTGATGACATG 356

QY 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
DB 357 CAACCTTTCAAAGCAGAGCTCCCGAGAGGCACTGCTGCTCGAGCTGCCATTGAAATG 416

QY 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlnThrLeuArgAlaAsp 140
DB 417 CAAGAAAGTCTGTTCACAAATAAGAGCTTGGAAATGAAGAATGAACAAACATTGAGAGCAGAT 476

QY 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
DB 477 GAGATATCTCCATCAGATCCAAACAAAGAGACTATGAGAAAGTTCTTGGGATTCTGAG 536

QY 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
DB 537 AGTCTCTGTGAGAGCTGTTTACAGAGGATGTGTGTGTACCAAGGCTACACATCAAAAA 596

QY 181 GluIleAspLysIleAsnGlyLysLeuGluClySerProAspAsnAspGlyPheLeuLys 200
DB 597 GAAATAGATAAAATAAATGGAAAATTAAGAGAGTCTCTCTGATAATGATGGTTTCTGAAG 656

QY 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
DB 657 GCTCCCTGCAGAAATGAAGTTTCTATTCCAACATAAGCCTTAGAATTTGATGAGCATGCAA 716

QY 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
DB 717 ACTTTCAAAGCAGAGCTCCCGAGAGGCACTGCTGCTTCGAGGCTGCCATTGAAATGCAA 776

QY 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGlnThrLeuArgAlaAspGln 260
DB 777 AAGTCTGTTCCAAATAAGAGCTTGGAAATGAAGAATGAACAAACATTGAGAGCAGATCAG 836

QY 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
DB 837 ATGTTCCCTTCAGATCAAAACAAAGAAACGTTGAAGAAATTTCTGGGATTCTGAGAGT 896

QY 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
DB 897 CTCGCTGAGACTGTTTACAGAGAGGATGTGTGTGTACCCAGGCTACACATCAAAAAAGAA 956

QY 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
DB 957 ATGGATAAAATAAGTGGAAAATTAGAAGATTCAACTAGCCTATCAAAAATCTTGGATACA 1016

QY 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
DB 1017 ATTCATTTCTGTGAAGAGGCAAGGAACTTCAAAAGATCACTGTGAACAATGTACAGGA 1076

QY 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
DB 1077 AAAATGGAAACAAATGAAAAGAAAGTTTGTGTGTACTGAAAAGAAAGAACTGTGAGAAGCAAAA 1136

QY 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysValLysTrpGluGlnLeuCysSerVal 380
DB 1137 GAAATAAATTCAGATTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 1196

QY 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
DB 1197 AGGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 1250

RESULT 2
AAAS9014
ID AAAS9014 standard; cDNA; 2683 BP.
XX AC AAAS9014;
XX DX 07-NOV-2000 (first entry)
XX DE Nucleotide sequence of BS322 full length cDNA sequence.
XX KW BS322; breast tissue marker; breast disease; breast cancer; ss.
XX OS Homo sapiens.
XX PN WO200043420-A1.
XX PD 27-JUL-2000.
XX PF 21-JAN-2000; 2000WO-US001452.
XX PR 21-JAN-1999; 99US-00234716.
XX PA (ABBO) ABBOTT LAB.
XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
XX PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
XX PI Stroupe SD;
XX DR WPI; 2000-499217/44.
XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
PT detection and diagnosis of breast disease especially breast cancer.
XX

Claim 12; Page 118-119; 126pp; English.

PS The present sequence represents the full length cDNA sequence encoding
 XX BS322 polypeptide. BS322 is a breast tissue marker. The BS322
 CC polynucleotides and polypeptides are used to detect and diagnose breast
 CC diseases, e.g. breast cancer. The BS322 polynucleotides are useful as a
 CC source of probes and primers, and the BS322 polypeptides are useful as
 CC antigens

XX Sequence 2683 BP; 1043 A; 462 C; 502 G; 676 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.94e-176 Length: 2683
 Score: 2052.00 Matches: 398
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-489-079-24 (1-398) x AAAS59014 (1-2683)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 57 ATGCATAAAGTCTGTTCCAAATAAAGCCTTGGAAATGGAAGAATGAACAAACATTGAGACGA 116
 Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTriaAspSer 40
 Db 117 GATGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCT 176
 Qy 41 GluSerLeuLysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
 Db 177 GAGAGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTACCCAGGCTCGCATCAA 236
 Qy 61 LysGluLeuLeuAspLysLeuGlnLysGlnLysSerProValLysAspGlyLeuLeu 80
 Db 237 AAAGAAATAGATAAAATAAATGGAATAATGGAAGGCTCTCTGTTAAAGATGGTCTTCTG 296
 Qy 81 LysAlaAsnCysGlyMetLysValSerLysAlaLeuGluLeuMetAspMet 100
 Db 297 AAGGCTAACTCGGGAATGAAAGTTCTATTCCAACTAAGGCTTAGAATTGATGACATG 356
 Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaLysGluMet 120
 Db 357 CAACCTTTCAAAGCAGAGCCTCCCGAAGGACCATCTGCCTTCGAGCCTGCCATTGAAATG 416
 Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
 Db 417 CAAAAGTCTGTTCCAAATAAAGCCTTGGAAATGGAAGATGAACAAACATTGAGACGAT 476
 Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTriaAspSerGlu 160
 Db 477 GAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAG 536
 Qy 161 SerLeuLysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaThrHisGlnLys 180
 Db 537 AGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTACCCAGGCTACACATCAAAA 596
 Qy 181 GluLeuLeuAspLysLeuGlnLysGlnLysSerProAspAsnAspGlyPheLeuLys 200
 Db 597 GAAATAGATAAATAAATGGAATAATGGAAGAGTCTCTGTAATGAATGATGTTTCTGAG 656
 Qy 201 AlaProCysArgMetLysValSerLysAlaLeuGluLeuMetAspMetGln 220
 Db 657 GTCCTCTGCAAGATGAAAGTTCTATTCCAACTAAGGCTTAGAATTGATGACATGCNA 716
 Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaLysGluMetGln 240
 Db 717 ACTTTCAAAGCAGAGCCTCCCGAAGGACCATCTGCCTTCGAGCCTGCCATTGAAATGCAA 776
 Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
 Db 777 AAGTCTGTTCCAAATAAAGCCTTGGAAATGGAAGATGAACAAACATTGAGACGATCAG 836

Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluLysAsnSerTriaAspSerGluSer 280
 Db 837 ATGTTCCCTTTCAGAAATCAAAACAAAGACGTTTGAAGAAAATTTCTGGGATTTCTGAGAGT 896
 Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
 Db 897 CTCGTCGAGACTGTTTTCAGAGAAGATGTGTGTGTACCAAGGCTACACATCAAAAGAA 956
 Qy 301 MetAspLysLysSerGlyLysLeuGluAspSerThrSerLeuSerLysLysLeuAspThr 320
 Db 957 ATGGATAAATAAGTGGAAAATTAAGAAGATTCACTAGCTTATCAAAAATCTTGATACA 1016
 Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
 Db 1017 ATTCATTTCTTGTGAAGAGCAAGGAACTTCAAAAAGATCACTGTGAACAAATGTACAGGA 1076
 Qy 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
 Db 1077 AAAATGGAACAAATGAAAAGAGTTTGTGTACTGAAAAGAACTGTGAGAGCAAAA 1136
 Qy 361 GluLeuLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
 Db 1137 GAATAAATCACAAGTTAGAGAACCAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 1196
 Qy 381 ArgPheLeuThrLeuMetLysMetLysLysLysLysLysLysLysLysLysLysLys 398
 Db 1197 AGGTTTCTCACACTCATGAAAATGAAAATTTATCTTTACATGAAAATTTGCATGT 1250

RESULT 3
 AAS47405
 ID AAS47405 standard; cDNA; 3681 BP.
 XX AAS47405;
 XX
 XX 18-DEC-2001 (first entry)
 XX
 DE Human cDNA cloneB726P-spliced_seq_B726P encoding a breast cancer protein.
 XX
 KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200179286-A2.
 XX
 XX 25-OCT-2001.
 PD
 XX
 PF 12-APR-2001; 2001WO-US012164.
 XX
 PR 17-APR-2000; 2000US-00551621.
 PR 08-JUN-2000; 2000US-00590751.
 PR 22-JUN-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 XX
 DR WPI: 2001-611721/70.
 DR P-PSDB; AAU33346.
 XX
 XX Breast Tumor Proteins and nucleic acids useful for the prevention,
 PT diagnosis and treatment of breast cancer.
 XX
 PS Claim 1; Page 271-272; 297pp; English.
 XX
 CC The invention relates to isolated breast tumour proteins and nucleic
 CC acids that encode them, including immunogenic fragments of the proteins.
 CC Also included are expression vectors expressing the proteins, transformed
 CC cells and antibodies raised against the proteins or an antigen presenting
 CC cell expressing the protein. The proteins and nucleic acids may be used
 CC in the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate breast tumour protein expression, i.e. breast tumours and

CC breast cancer e.g by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a cDNA from a
CC breast tumour cDNA library isolated by subtractive hybridisation against
CC a normal breast cDNA library and encodes a breast tumour protein of the
CC invention. The present sequence is also a splice variant
XX

SQ Sequence 3681 BP; 1412 A; 680 C; 758 G; 826 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.: 1e-173 Length: 3681
Score: 2024.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.64% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-24 (1-398) x AAS47405 (1-3681)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1308 ATGCATAAAGTCTGCCCAATAAAGCCCTGGAAATTGAAAAATGAACAACATTCGAGACCA 1367
Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1368 GATGAGATACCTCCCATCAGAAATCAACAAAGAGACTATGAAGAAAGTTCTTGGGATCTCT 1427
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGAGACTGTTTTCACAGAGGATGTGTGTATCCCAAGGCTRCRCATCAA 1487
Qy 61 LysGluLeuAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1488 AAGAAATAGATAAAATAAATGGAATAATGAAGGGTCTCTGTTAAAGATGGTCTTCTG 1547
Qy 81 LysAlaAsnCysGlyMetLysValSerLysIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGCTTAACTCGGAATGAAGTTTCTATTCACACTAAAGCCTTAGAATTTGATGGACATG 1607
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1608 CAACCTTCAAGCAGAGCCCTCCGAGAGCCATCTGCCCTCGAGCCTGCCATTTGAAATG 1667
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1668 CAAAGTCTGTTCCTCAATAAAGCCCTTGGAAATTGAAGAAATGAACAACATTCGAGACGAGAT 1727
Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1728 GAGATACCTCCCATCAGAAATCAACAAAGAGACTATGAAGAAAGTTCTTGGGATCTCGAG 1787
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGTGAGACTGTTTTCACAGAAGATGTGTGTATCCCAAGGCTRCRCATCAAAAA 1847
Qy 181 GluLeuAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 1848 GAAATAGATAAAATAAATGGAATAATTAAGAGAGTCTCTCTATATGATGTTCTTCTGAAG 1907
Qy 201 AlaProCysArgMetLysValSerLysIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1908 GCTCCCTGCGAAGTGAAGTTTCTATTCACACTAAAGCCTTAGAATTTGATGGACATGCAA 1967
Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240

Db 1968 ACTTTCAAAGCAGAGCCCTCCGAGAGCCATCTGCCCTTCGAGCTGCCATTGAAATGCAA 2027
Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
Db 2028 AAGTCTGTTCCTCAATAAAGCCTTGGAAATTGAAGAAATGAACAACATTCGAGACGATCAG 2087
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 2088 ATGTTCCCTTCAGAAATCAAAACAAAGAAAGTGAAGAAATTTCTTGGGATTTCTGAGAGT 2147
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 2148 CTCGTGAGACTGTTTCACAGAAGGATGTGTGTGTACCCCAAGGCTACACATCAAAAAGAA 2207
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 2208 ATGCATAAATAAAGTGGAAATTTAGAAAGATTCAACTAGCCCTATCAAAAATCTTGGATACA 2267
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 2268 GTTCATTTCTGTGAAGAGCAGAGGAACTTCAAAAAGATCCTGTGAAACACGTCACAGGA 2327
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2328 AAAATGGAACAAATGAAAAAGAAAGTTTGTGTGTACTGAAAAAGAAACTGTTCAGAAGCAAAA 2387
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysValLysTrpGluGlnLeuCysSerVal 380
Db 2388 GAAATAAATTCACAGTTTAGAGAACCAAAAAGTTAAATGGGAACAGAGCTCTCGAGTGTG 2447
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
Db 2448 AGGTTTCTCACTCATGAAANAATGAAANAATATCTTTACATGAAANAATTGCATGT 2501

RESULT 4

AAS47422
ID AAS47422 standard; cDNA; 2232 BP.

XX AAS47422;

XX 18-DEC-2001 (first entry)

XX Human cDNA encoding breast cancer protein B726P fusion protein #2.

XX Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX gene therapy.

XX Homo sapiens.

XX WO200179286-A2.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US012164.

XX 17-APR-2000; 2000US-00551621.

XX 08-JUN-2000; 2000US-00590751.

XX 22-JUN-2000; 2000US-00604287.

XX 20-JUL-2000; 2000US-00620405.

XX (CORI-) CORIXA CORP.

XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;

XX WPI; 2001-611721/70.

XX P-PSDB; AAU33358.

XX Breast Tumor Proteins and nucleic acids useful for the prevention,
XX diagnosis and treatment of breast cancer.

XX Claim 37; Page 291-292; 297pp; English.

XX The invention relates to isolated breast tumour proteins and nucleic

acids that encode them, including immunogenic fragments of the proteins.
 CC Also included are expression vectors expressing the proteins, transformed
 CC cells and antibodies raised against the proteins or an antigen presenting
 CC cell expressing the protein. The proteins and nucleic acids may be used
 CC in the prevention, diagnosis and treatment of diseases associated with
 CC inappropriate breast tumour protein expression, i.e. breast tumours and
 CC breast cancer e.g by gene therapy. The nucleic acids and their
 CC complements may also be used as DNA probes in diagnostic assays to detect
 CC and quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. The
 CC proteins, nucleic acids and antibodies may be used in assays to identify
 CC modulators (e.g. antagonists) of breast tumour protein expression and
 CC activity. The antibodies and antagonists may also be used to down
 CC regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the proteins in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)) and in other immuno-
 CC purification diagnostic techniques. The present sequence is a cDNA from a
 CC breast tumour cDNA library isolated by subtractive hybridisation against
 CC a normal breast cDNA library and encodes a breast tumour protein of the
 CC invention

XX Sequence 2232 BP; 826 A; 399 C; 467 G; 535 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.: 6.63e-174 Length: 2232
 Score: 2023.00 Matches: 393
 Percent Similarity: 98.99% Conservative: 1
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.59% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-24 (1-398) x AAS47422 (1-2232)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 1036 ATGCAAAAGTCTGCCCAATAAAGCCCTGGAAATTTGAAATAATGAACAAACATTGAGAGCA 1095
 Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 1096 GATGAGTACTCCCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTTCT 1155
 Qy 41 GluSerLeuLysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
 Db 1156 GAGAGTCTCTGTGAGACTGTTTACAGAGGATGTGTGTTTACCCCAAGGCTRCRCATCAA 1215
 Qy 61 LysGluLeuAspLysLeuGlnLysGluGlySerProValLysAspGlyLeuLeu 80
 Db 1216 AAAGAAATAGATAAAATAAATGCAAAATTTAGAAAGGCTCTCCTGTTTAAAGATGCTTCTG 1275
 Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
 Db 1276 AAGGCTAAGTCCGGAATGAAGTTTCTATTCACAACTTAAGCCCTTAGAATTGATGACATG 1335
 Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
 Db 1336 CAAACTTTCAAAGCAGAGCCTCCGAGAGGCACTCTGCCTTCGAGCCTGCCATTGAAATG 1395
 Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
 Db 1396 CAAAAGTCTGTTCCAAATAAAGCCTTGGAAATTTGAAGAAATGAACAAACATTGAGACAGAT 1455
 Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
 Db 1456 GAGATACTCCCATCAGATCCCAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGNG 1515
 Qy 161 SerLysGluGluThrValSerGlnLysAspValCysLeuProLysAlaAlaThrHisGlnLys 180
 Db 1516 AGTCTCTGTGAGACTGTTTACAGAGGATGTGTGTTTACCCCAAGGCTRCRCATCAAAA 1575
 Qy 181 GluLeuAspLysLeuGlnLysGluGlySerProAspAsnAspGlyPheLeuLys 200
 Db 1576 GAAATAGATAAAATAAATGAAGAAATTTAGAAAGAGTCTCCTGATGAATGATGTTTCTGAAG 1635

Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
 Db 1636 GCTCCCTGCAGAAATGAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGCATGCAA 1695
 Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
 Db 1696 ACTTTCAAAGCAGAGCCTCCGAGAGCCATCTGCCCTTCGAGCCTGCCATTGAAATGCAA 1755
 Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
 Db 1756 AAGTCTGTTCCAAATAAAGCCTTGGAAATTTGAAGAAATGAACAAACATTGAGAGCAGATCAG 1815
 Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluLeuAsnSerTrpAspSerGluSer 280
 Db 1816 ATGTTCCCTTCAGAAATCAAAACAAAGAAAGCTTGAAGAAATTTCTTGGGATTTCTGAGAGT 1875
 Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
 Db 1876 CTCGTTGAGACTGTTTACAGAGAGATGTGTGTGTACCCCAAGGCTACACATCNAAGAA 1935
 Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLysSerLysIleLeuAspThr 320
 Db 1936 ATGATATAAATAAGTGGAAATTTAGAGATTTCAACTAGCCTATCAAAAATCTTGATACA 1995
 Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
 Db 1996 GTTCATTCTTGTGAAGAGAGCAAGGGAACCTTCAAAAAGATCACTGTGAAACAACGTACAGGA 2055
 Qy 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
 Db 2056 AAAATGGAAACAAATGAAAGAAAGTGTGTGTACTGAAAGAAAGAACTGTGAGAGCAAAA 2115
 Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
 Db 2116 GAATATAAATCACAGTTAGAGAACCAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2175
 Qy 381 ArgPheLeuThrLeuMetLysMetLysIleLeuSerTyrMetLysIleAlaCys 398
 Db 2176 AGGTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATAATTTGCAATG 2229

RESULT 5
 ABS64023 standard; DNA; 2232 BP.
 ID ABS64023
 XX ABS64023;
 AC ABS64023;
 XX
 DT 15-NOV-2002 (first entry)
 DE Human breast tumour polynucleotide #472.
 XX
 DE Human; breast tumour protein; gene; ds; breast cancer; cytostatic;
 KW vaccine.
 KW
 XX Homo sapiens.
 OS
 XX
 PN US2002085998-A1.
 XX
 XX 04-JUL-2002.
 PD
 XX
 PF 13-APR-2001; 2001US-00834759.
 XX
 XX 28-DEC-1998; 98US-00222575.
 PR 02-APR-1999; 99US-00285480.
 PR 23-JUN-1999; 99US-00339338.
 PR 02-SEP-1999; 99US-00389681.
 PR 03-NOV-1999; 99US-00433826.
 PR 17-APR-2000; 2000US-00551621.
 PR 08-JUN-2000; 2000US-00590751.
 PR 22-JUN-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 XX
 XX (CORI-) CORIXA CORP.
 XX

PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 PI Henderson RA;
 XX WPI; 2002-635657/68.
 DR P-PSDB; ABG78925.
 XX Novel breast cancer polynucleotides and polypeptides encoded by the
 PT polynucleotides, useful for detecting the presence of breast cancer in a
 PT patient, and in pharmaceutical compositions, for treating breast cancer.
 XX Disclosure; Page 221-222; 247pp; English.
 XX The invention relates to an isolated breast tumour polynucleotide and the
 CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
 CC detecting the presence of breast cancer in a patient, and in
 CC pharmaceutical compositions for treating breast cancer. The sequences are
 CC useful for stimulating an immune response in a patient and can therefore
 CC be used in production of vaccines. The sequences are also useful for
 CC detecting the presence of a cancer in a patient, by obtaining a
 CC biological sample from the patient, contacting the biological sample with
 CC a composition of the invention and detecting the amount of polynucleotide
 CC that hybridizes to the sample. This sequence represents a human breast
 CC tumour polynucleotide of the invention
 XX
 SQ Sequence 2232 BP; 826 A; 399 C; 467 G; 535 T; 0 U; 5 Other;
 Alignment Scores:
 Pred. No.: 6.63e-174 Length: 2232
 Score: 2023.00 Matches: 393
 Percent Similarity: 98.99% Conservative: 1
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.59% Indels: 0
 DB: 6 Gaps: 0
 US-09-489-079-24 (1-398) x ABS64023 (1-2232)
 QY 1 MetGlnLysSerValProAsnLysAlaLeuLeuLysAsnGluGlnThrLeuArgAla 20
 DB 1036 ATGCATAAAGTCTGCTCCCAATTAAGCTTTGGAATTTGAATAATGACAAATTTGAGAGCA 1095
 QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 DB 1096 GATGAGATACTCCCATCAGAAATCAAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 1155
 QY 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
 DB 1156 GAGAGTCTCTGTGAGACTGTGTTACAGAGGATGTGTGTTTACCAAGGCTRCRCATCAA 1215
 QY 61 LysGluLeuAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
 DB 1216 AAGAAATAGATAAATAATTAATGGAATTTAGAGGGTCTCTGTTAAAGATGGTCTTCTG 1275
 QY 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
 DB 1276 AAGGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTTAGTGACATG 1335
 QY 101 GlnThrPhelysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120
 DB 1336 CAAACTTTCGAAGCAGAGCTCCCGAAGGCCATCTGCCCTTCGAGCCTGCCATTGAAATG 1395
 QY 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
 DB 1396 CAAAGTCTGTTCCAAATTAAGCTTTGGAATTTGAAGATGACAAACATTTGAGAGCAGAT 1455
 QY 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
 DB 1456 GAGATACTCCCATCAGAAATCAAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGAG 1515
 QY 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
 DB 1516 AGTCTCTGTGAGACTGTGTTACAGAGGATGTGTGTTTACCAAGGCTRCRCATCAAAA 1575
 QY 181 GluIleAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200

DB 1576 GAAATAGATAAATAAATGGAATAATTAGAAGACTCTCTGATAATGATGTTTCTGAG 1635
 QY 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
 DB 1636 GCTCCCTCGCAGAAATGAAAGTTTCTATTCCAACTAAAGCCCTTAGAATTTGATGSCATGCAA 1695
 QY 221 ThrPhelysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
 DB 1696 ACTTTCAAGCAGAGCTCCCGAAGGCCATCTGCCCTTCGAGCCTGCCATTGGAATGCAA 1755
 QY 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
 DB 1756 AAGTCTGTTCCAAATAAAGCCTTTGGAATTTGAAGAATGAACAAACATTTGAGAGCAGATCAG 1815
 QY 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
 DB 1816 ATGTTCCCTTCAGAAATCAAAACAAAGAGTTGGAAGAAATTTCTTGGGATTTCTGAGAGT 1875
 QY 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
 DB 1876 CTCGGTGAGACTGTTTCACAGAGGATGTGTGTACCCAGGCTACACATCAAAAGAA 1935
 QY 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
 DB 1936 ATGCATAAATAAAGTGGAAATTTAGAAGATTCAACTAGCCTATCAAAATCTTTGGATACA 1995
 QY 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
 DB 1996 GTTCATTCTTTGTGAAAGAGCAGAGGAACTTTCAAAAAGATCCTGTGAACAACTGACAGGA 2055
 QY 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
 DB 2056 AAAATGGACAAATGAAGAAGATTTTGTGTACTGAAAAGAAACTGTCTCAGAGGCAAAA 2115
 QY 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysValLysTrpGluGlnGluLeuCysSerVal 380
 DB 2116 GAAATAAATCAGCTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2175
 QY 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
 DB 2176 AGTTTTCTCACACTCATGAAATGAAATTTATCTCTTACATGAAATTTGCATGT 2229
 RESULT 6
 ABT33235
 ID ABT33235 standard; DNA; 2232 BP.
 XX
 AC ABT33235;
 XX
 DT 15-MAY-2003 (first entry)
 XX
 DE Human tumour-related DNA sequence - SEQ ID No 491.
 XX
 KW Human; ds; vaccine; gene therapy; T cell stimulation; T cell expansion;
 KW tumour; breast cancer; cancer; immune response stimulation.
 XX
 OS Homo sapiens.
 XX
 PN W020283956-A1.
 XX
 PD 24-OCT-2002.
 XX
 PF 15-APR-2002; 2002WO-US012378.
 XX
 PR 13-APR-2001; 2001US-00834759.
 PR 07-DEC-2001; 2001US-00007805.
 PR 13-FEB-2002; 2002US-00076622.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
 PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
 PI Vedvick TS, Mcneill PD, Durham M;

XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX
XX
PS Disclosure; Page 308-309; 375pp; English.
XX
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumor protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumors (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumor protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present DNA sequence represents a
CC human tumor-related DNA sequence
XX
SQ Sequence 2232 BP; 826 A; 399 C; 467 G; 535 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 6,63e-174 Length: 2232
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 10 Gaps: 0

US-09-489-079-24 (1-398) x ABT33235 (1-2232)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1036 ATGCAAAAGTCTGCCAAATAAGCCTTGGAAATGAAATAATGAACAAACATTGAGACA 1095

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1096 GATGAGACTCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCT 1155

QY 41 GluSerLeuGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1156 GAGAGTCTCTGTGAGACTGTTTACAGAGAGGATGTTGTTTACCCAGGCTTCRCATCAA 1215

QY 61 LysGluLeuAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1216 AAAGAAATAGATAAATAAATGAAATGAAAGGCTCTCTGTTAAAGATGGTCTTCTG 1275

QY 81 LysAlaAsnCysGlyMetLysValSerLeuProThrLysAlaLeuGluLeuMetAspMet 100
Db 1276 AAGGCTAACTGCGGAATGAAGATTTCTATTCCAACCTAAAGCCTTGAATTTGATGACATG 1335

QY 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1336 CAATCTTTCAAGCAGAGAGCTCCGAGAGAGCCATCTGCCTTCGAGCCTGCCATTGAAATG 1395

QY 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1396 CAAAAGTCTGTTCCAAATAAAGCCTTGGAAATGAAAGATGAACAAACATTGAGAGCAGAT 1455

QY 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1456 GAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAG 1515

QY 161 SerLeuGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1516 AGTCTCTGTGAGACTGTTTACAGAGAGGATGTTGTTTACCCAGGCTTCRCATCAAATA 1575

QY 181 GluLeuAspLysIleAsnGlyLysLeuGluGlySerProAspAsnAspGlyPheLeuLys 200
Db 1576 GAAATAGATAAATAATGAAATGAAAGGCTCTCTGATGATGATGATGATGATGATGATG 1635

QY 201 AlaProCysArgMetLysValSerLeuProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1636 GCTCCCTGCAGATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGACATGCAA 1695

221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
1696 ACTTTCAAAGCAGAGCCTCCCGAGAGCCATCTGCTTCGAGCCTGCCATTGAATGAAATGCAA 1755

241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
1756 AAGTCTGTTCCAAATAAAGCCTTGGAAATGAAAGATGAACAAACATTGAGAGCAGATCAG 1815

261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
1816 ATGTTCCCTTCAGAAATCAAAACAAAGAGAGCTTGAAGAAATTTCTTGGATTCTGAGAGT 1875

281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
1876 CTCGTCGAGACTGTTTTCACAGAGAGATGTTGTGTGTACCCAAAGGCTACATCATCAAAAGAA 1935

301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
1936 ATGATATAAATAAGTGGAAATTTAGAAGATTTCAATAGCCTATCAAAAATCTTGGATACA 1995

321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
1996 GTTCATTTCTTGTGAAGAGCAAGGGAATCTTCAAAAAGATCACTGTGAACAACAGTACAGGA 2055

341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
2056 AAAATGGAACAAATGAAAGAGTGTGTGTGTACTGAAAAAGAACTGTCTCAGAGCAAAA 2115

361 GluLeuLysSerGlnLeuGluAsnGlnLysValLysValLysValLysValLysValLys 380
2116 GAAATAAATAACACAGTTAGAGAACCAAAAGTTAAATGGGAACAGAGCTCTGCAAGTGTG 2175

381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
2176 AGTTTCTCACACTCATGAAATGAAATTTATCTTTACATGAAATTTGCAATG 2229

RESULT 7
ADL93153
ID ADL93153 standard; DNA; 2232 BP.
XX AC ADL93153;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human breast cancer-associated polypeptide fusion protein DNA #2.
XX
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human; ds;
KW gene.
XX
OS Homo sapiens.
XX
PN US2003166022-A1.
XX
PD 04-SEP-2003.
XX
PF 15-APR-2002; 2002US-00124805.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-0051621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX

PI Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.
DR P-PSDB; ADL93156.
XX
PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX
XX Disclosure; SEQ ID NO 491; 294pp; English.
XX
CC The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents DNA encoding a human breast cancer-associated
CC polypeptide fusion protein.
XX
SQ Sequence 2232 BP; 826 A; 399 C; 467 G; 535 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 6.63e-174 Length: 2232
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 11 Gaps: 0

US-09-489-079-24 (1-398) x ADL93153 (1-2232)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAla 20
Db 1036 ATGCAAGCTCTGCTCCCAATTAAGCCTTGGAAATGAAATGAACAAATGAGACGA 1095
Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1096 GATGAGATACCTCCCATCAGATCCAAACAAAGAGACTATGAAGAAAGTTCTTGGGATTC 1155
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1156 GAGAGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTACCAAGGCTRCRCATCAA 1215
Qy 61 LysGluLeuAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1216 AAAGAAATAGATAAATAAATGAAGATTTAGAGGGTCTCTGTTAAAGATGGTCTTCG 1275
Qy 81 LysAlaLysCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 1276 AAGCTTAACTCGGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATG 1335
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1336 CAAACTTTCAAGCAGAGCTCCCGAGAGCCATCTGCCTTCGAGCCTGCATTTGAATG 1395
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAsp 140
Db 1396 CAAAGTCTGTTCCTCAATAAAGCTTGGAAATGAAGATGAACAAACATTGAGACAGAT 1455
Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1456 GAGATACCTCCCATCAGAAATCCAAACAAAGAGACTATGAAGAAAGTTCTTGGGATTTCTGAG 1515
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1516 AGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTACCAAGGCTRCRCATCAAAAA 1575
Qy 181 GluLeuAspLysIleAsnGlyLysLeuGluGlySerProAspAsnAspGlyPheLeuLys 200
Db 1576 GAATAGATAAATAAATGAAGATTTAGAGAGTCTCTGATATGATGTTTCTCAAG 1635
Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1636 GCTCCCTGCAGAAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATGCA 1695

Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1696 ACTTTCAAGCAGAGCTCCCGAGAGCCATCTGCCTTCGAGCTGCCATTGAAATGCAA 1755
Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGln 260
Db 1756 AAGTCTGTTCCAAATAAAGCTTGGAAATGAAGATGAACAAACATTGAGACAGATCAG 1815
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 1816 ATGTTCCCTTCAGAAATCAAAACAAAGAGTTGAAGAAATTTCTGGGATTTCTGAGAGT 1875
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 1876 CTCGTGAGACTGTTTCAAGAGAGATGTGTGTACCCCAAGGCTACACATCAAAAGAA 1935
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 1936 ATGGATAAATAAAGTGGAAATTAGAAGATTCAATAGCCTATCAAAATCTTTGGATACA 1995
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAsnHisCysGluGlnCysThrGly 340
Db 1996 GTTCATCTTGTGAAAGAGCAGGGAACCTTCAAAAGATCCTGTGAACCAACGTACAGGA 2055
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2056 ARAATGGACAAATGAAAGAGATTTTGTGTACTGAAAAAGAACTGTTCAGAAAGCAAAA 2115
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 2116 GAAATAAATACAGCTTAGAGAACCAAAAGTTAAATGGGAACAGAGCTCTCGAGTGTG 2175
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleLysSerTyrMetLysIleAlaCys 398
Db 2176 AGGTTTCTCACTCATGAAATGAAATATTCTCTTACATGAAATTTGCATGT 2229

RESULT 8
AAF17975
ID AAF17975 standard; DNA; 3681 BP.
XX AAF17975;
AC AAF17975;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human breast cancer associated B726P consensus sequence #1.
XX
KW Human; breast cancer associated gene; vaccine; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200060076-A2.
XX
PD 12-OCT-2000.
XX
PF 15-FEB-2000; 2000WO-US005308.
XX
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
XX
PA (CORI-) CORIXA CORP.
XX
PI Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX
DR WPI; 2001-122627/13.
XX P-PSDB; AAB50245.
XX
PT An isolated polypeptide useful for the treatment and diagnosis of tumors
PT e.g. breast cancer comprises at least an immunogenic portion of a breast
XX tumor protein.
PS Claim 6; Page 225-227; 238pp; English.

XX The present invention provides the coding sequences and some protein
CC sequences of proteins associated with breast cancer in humans. These
CC sequences can be used in the diagnosis and treatment of cancers,
CC particularly breast tumors
XX

SQ Sequence 3681 BP; 1411 A; 680 C; 758 G; 826 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 1,23e-173 Length: 3681
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.53% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-24 (1-398) x AAF17975 (1-3681)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 1308 ATGCAAAAGTCTGCCAAATAAAGCCTTGGAAATTGAAAAATGAACAAACATTGAGAGCA 1367

Qy 21 AspGluLeuLeuProSerGlnSerLysGlnLysAspTyrGluGluSerSerTpaAspSer 40
Db 1368 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 1427

Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGAGACTGTTTACAGAGAGGATGTTGTTTACCCAGGCTTCRCATCAA 1487

Qy 61 LysGluLeuAspLysAlaAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1488 AAAGAAATAGATAAAATAAATGAAATAGAAAGGCTCTCCTGTTAAAGATGGTCTTCTG 1547

Qy 81 LysAlaAsnCysGlyMetLysValSerLeuProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGGCTAACTCGGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTGATGACATG 1607

Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1608 CAACTTTCAAGCAGAGCCTCCGAGAGGCATCTGCTTCGAGCCTGCCATTGAATG 1667

Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1668 CAAAAGTCTGTTCCAAATAAAGCCTTGGAAATTGAAGAAATGAACAAACATTGAGAGCAGAT 1727

Qy 141 GluLeuLeuProSerGlnSerLysGlnLysAspTyrGluGluSerSerTpaAspSerGlu 160
Db 1728 GAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCTGAG 1787

Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGAGACTGTTTACAGAGAGGATGTTTATCCCACTAAAGCCTTAGAATTGATGACATGCAA 1847

Qy 181 GluLeuAspLysAlaAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 1848 GAAATAGATAAATAAATGAAATAGAAAGTCTCCTGATAATGATGTTTCTGAG 1907

Qy 201 AlaProCysArgMetLysValSerLeuProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1908 GCTCCCTGAGAAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTGATGACATGCAA 1967

Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1968 ACTTTCAAAGCAGAGCCTCCGAGAGGCATCTGCTTCGAGCCTGCCATTGAAATGCAA 2027

Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
Db 2028 AAGTCTGTTCCAAATAAAGCCTTGGAAATTGAAGAAATGAACAAACATTGAGAGCAGATCAG 2087

Qy 261 MetPheProSerGlnSerLysGlnLysAsnValGluGluAsnSerTpaAspSerGluSer 280
Db 2088 ATGTTCCCTTCAGAAATCAAAACAAAGAGAGTTGAAGAAATTTCTTGGGATTCTGAGAGT 2147

Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 2148 CTCGCTGAGACTGTTTACAGAGAGATGTTGTGTAGCCCAAGGCTACACATCAAAAAGAA 2207

Qy 301 MetAspLysLysSerGlyLysLeuGluAspSerThrSerLeuSerLysAlaLeuAspThr 320
Db 2208 ATGGATAAAATAAGTGGAAAAATTAGAAGATTCAACTAGCCCTATCAAAAATTTGGATACA 2267

Qy 321 IleHisSerCysGluArgAlaAaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 2268 GTTCATTCTTGTGAAGAGCAAGGGAACCTTCAAAAGATCACTGTGAACAACGTACAGGA 2327

Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2328 AAAATGGAAACAAATCAAAAAGAGTGTGTTGTACTGAAAAAGAACTGTCTAGAAGCAAAA 2387

Qy 361 GluLeuLysSerGlnLeuGluAsnGlnLysValLysValLysValLysValLysValLysVal 380
Db 2388 GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAAACAGAGCTCTGCAGTGTG 2447

Qy 381 ArgPheLeuThrLeuMetLysMetLysLysLysLysLysLysLysLysLysLysLysLys 398
Db 2448 AGTTTCTCACACTCATGAAATGAAATATCTCTTACATGAAATTTGCATGT 2501

RESULT 9
AAI67218
ID AAI67218 standard; cDNA; 3681 BP.
XX AAI67218;
AC AAI67218;
XX
DT 11-FEB-2002 (first entry)
XX
DE B726P splice variant encoding cDNA.
XX
KW Genetic subtraction; DNA microarray analysis; polymerase chain reaction;
KW cancer; B726P; ss.
XX
OS Homo sapiens.
XX
PN WO200175171-A2.
XX
PD 11-OCT-2001.
XX
PF 02-APR-2001; 2001WO-US010631.
XX
PR 03-APR-2000; 2000US-0194241P.
PR 20-JUL-2000; 2000US-0219862P.
PR 27-JUL-2000; 2000US-0221300P.
PR 18-DEC-2000; 2000US-02256592P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Dillon DC, Molesch DA, Xu J, Zehentner B, Persing DH;
XX
XX WPI; 2001-626449/72.
XX P-PSDB; AAG65983.
XX
PT Identifying tissue (tumor)-specific polynucleotides overexpressed in
PT tissue of interest as compared to control tissue, for detecting cancer
PT cells in patient, comprises DNA microarray analysis or quantitative
PT polymerase chain reaction.
XX
PS Claim 4; Page 104-105; 127pp; English.
XX
CC The invention relates to identifying tissue-specific polynucleotides (P)
CC that involves performing a genetic subtraction to identify pool of (P)
CC from tissue of interest (TI), performing DNA microarray analysis to
CC identify first subset of polynucleotides (SPI) at least 2-fold over
CC expressed in TI, and performing quantitative polymerase chain reaction
CC (PCR) analysis on SPI to identify second subset of (P). The method is
CC useful for determining the presence or absence of a cancer cell in a
CC patient, monitoring the progression of cancer in a patient using a

CC biological sample such as blood, serum, lymph nodes, bone marrow, sputum,
 CC urine or a tumour biopsy sample. The methods are useful for determining
 CC the presence or absence of or monitoring progression of prostate, breast,
 CC colon, ovarian, lung, head and neck, lymphoma, leukemia, melanoma, liver,
 CC gastric, kidney, bladder, pancreatic or endometrial cancer. Sequences
 CC AA167218-223 represent determined cDNA sequences of splice variants of
 CC B726P

XX
 SQ Sequence 3681 BP; 1411 A; 680 C; 758 G; 826 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.: 1.23e-173 Length: 3681
 Score: 2023.00 Matches: 393
 Percent Similarity: 98.99% Conservative: 1
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.59% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-24 (1-398) x AA167218 (1-3681)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 1308 ATGCAAAAGTCTGTCCCAATAAAGCCTTGGAAATTGAAAATGAACAACATTGAGAGCA 1367
 Qy 21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 1368 GATGAGATATCTCCATCAGATCCAAACAAAGGACTATGAGNAAAGTTCCTGGGATTCT 1427
 Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
 Db 1428 GAGAGTCTCTGTGAGACTGTTTACAGAAAGATGTGTGTGTACCCAGGCTCRCAATCAA 1487
 Qy 61 LysGluLeuAspLysIleAsnGlyLysGluGlySerProValLysAspGlyLeuLeu 80
 Db 1488 AAGAAATAGATAAAATAAATGGAATTTAGAAAGGTCTCTGTGTAAGATGTCTTCTG 1547
 Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
 Db 1548 AAGCTTAACCTCCGGAATGAAGTTTCTATTCCAACTAAGCCTTAGAATTTGATGGACATG 1607
 Qy 101 GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120
 Db 1608 CAAACTTTCAAAGCAGAGCCTCCCGAGAAGCCATCTGCCCTTCGAGCCTGCCATTGAAATG 1667
 Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
 Db 1668 CAAAGTCTGTTCAAATAAAGCCTTGAATTTGAAGAAATGAACAACATTGAGAGCAGAT 1727
 Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
 Db 1728 GAGATATCTCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTTCCTGGGATTCTGAG 1787
 Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
 Db 1788 AGTCTCTGTGAGACTGTTTACAGAAAGATGTGTGTGTACCCAGGCTCRCAATCAAAA 1847
 Qy 181 GluLeuAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
 Db 1848 GAAATAGATAAAATAAATGGAATTTGAAGAGTCTCTGATATGATGTTTCTGAAG 1907
 Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
 Db 1908 GCTCCCTGCAGATGAAGTTTCTATTCCAACTAAGCCTTAGAATTTGATGGACATGCAA 1967
 Qy 221 ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
 Db 1968 ACITTCAAAGCAGAGCCTCCCGAGAAGCCATCTGCCCTTCGAGCCTGCCATTGAAATCAA 2027
 Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
 Db 2028 AAGTCTCTTCCAAATAAAGCCTTGGAAATGAAGAAATGAACAACATTGAGAGCAGATCAG 2087
 Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280

Db 2088 ATGTTCCCTTCAGAAATCAAAACAAAGASGTTGAAGAAATTTCTTGGGATTCTGAGAGT 2147
 Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
 Db 2148 CTCCTGAGACTGTTTTCACAGAAAGGATGTGTGTGTACCCCAAGGCTACACATCAAAAAGAA 2207
 Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
 Db 2208 ATGGATAAAATAAGTGGAAAATTAGAAAGATTCAACTAGCCTATCAAAAATCTTGGATACA 2267
 Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
 Db 2268 GTTCAATCTTGTGAAGAGCAGAGGAACCTTCAAAAAGATCACTGTGAACAACGTACAGGA 2327
 Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
 Db 2328 AAAATGGAACAAATGAAAAGAAAGTTTGTGTGTACTGAAAAAGAAACTGTTCAGAAGCAAAA 2387
 Qy 361 GluLeuLysSerGlnLeuGluAsnGlnLysValLysTTPGluGlnGluLeuCysSerVal 380
 Db 2388 GAAATTAATTCAGCTTAGAGAACCAAAAAGTTAAATGGGAACAGAGCTCTTCGAGTGTG 2447
 Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
 Db 2448 AGGTTTCTCACACTCATGAAATGAAAATTTATCTCTTACATGAAATTTGCATGT 2501

RESULT 10
 ABS64006
 ID ABS64006 standard; DNA; 3681 BP.
 XX
 AC ABS64006;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Human breast tumour polynucleotide #458.
 XX
 KW Human; breast tumour protein; gene; ds; breast cancer; cytostatic;
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN US2002085998-A1.
 XX
 PD 04-JUL-2002.
 XX
 PF 13-APR-2001; 2001US-00834759.
 XX
 PR 28-DEC-1998; 98US-00222575.
 PR 02-APR-1999; 99US-00285480.
 PR 23-JUN-1999; 99US-00339338.
 PR 02-SEP-1999; 99US-00389681.
 PR 03-NOV-1999; 99US-00433826.
 PR 17-APR-2000; 2000US-00551621.
 PR 08-JUN-2000; 2000US-00590751.
 PR 22-JUN-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 PI Henderson RA;
 PI
 DR WPI; 2002-635657/68.
 DR P-PSDB; ABG78913.
 XX
 PT Novel breast cancer polynucleotides and polypeptides encoded by the
 PT polynucleotides, useful for detecting the presence of breast cancer in a
 PT patient, and in pharmaceutical compositions, for treating breast cancer.
 XX
 PS Claim 1; Page 195-197; 247pp; English.
 XX
 CC The invention relates to an isolated breast tumour polynucleotide and the

CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
CC detecting the presence of breast cancer in a patient, and in
CC pharmaceutical compositions for treating breast cancer. The sequences are
CC useful for stimulating an immune response in a patient and can therefore
CC be used in production of vaccines. The sequences are also useful for
CC detecting the presence of a cancer in a patient, by obtaining a
CC biological sample from the patient, contacting the biological sample with
CC a composition of the invention and detecting the amount of polynucleotide
CC that hybridizes to the sample. This sequence represents a human breast
CC tumour polynucleotide of the invention
XX
SQ Sequence 3681 BP; 1411 A; 680 C; 758 G; 826 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 1,23e-173 Length: 3681
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-24 (1-398) x ABS64006 (1-3681)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAla 20
Db 1308 ATGCAAAAGTCTGTCCCAATAAAGCCTTGGAAATTTGAAAAATGAACAAACATTGAGAGCA 1367
Qy 21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1368 GATGAGATACCTCCCATCAGATCCCAACAAAGGACTATGAAGAAAGTTCTTTGGGATTC 1427
Qy 41 GluSerLeuGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGTGAGACTGTTTCACAGAGGATGTGTGTTTACCCAAAGCTTCRCATCAA 1487
Qy 61 LysGluLeuAspLysLeuGlnLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 1488 AAAGAAATAGATAAAATAAATGAAAAATTAGAAGGCTCTCCTGTTAAAGATGGTCTTC 1547
Qy 81 LysAlaAsnGlyMetLysValSerLeuProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGGCTAACTCGGGAATGAAAGTTTCTATTTCAACTAAAGCCCTTAGAATTTGATGACATG 1607
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaLeuMet 120
Db 1608 CAACTTTCAAGCAGAGCCTCCCGAAGCCTCTGCTTCGACCTGCCATTGAATG 1667
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAsp 140
Db 1668 CAAAAGTCTGTTTCCAAATAAAGCCTTGGAAATTTGAAGAAATGAACAAACATTGAGAGCAGAT 1727
Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1728 GAGATACCTCCCATCAGATCCCAACAAAGGACTATGAAGAAAGTTCTTTGGGATTCCTGAG 1787
Qy 161 SerLeuGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGTGAGACTGTTTCACAGAGGATGTGTGTTTACCCAAAGCTTCRCATCAAAA 1847
Qy 181 GluLeuAspLysLeuGlnLysLeuGluGlySerProAspAsnAspGlyPheLeuLys 200
Db 1848 GAAATAGATAAAATAAATGAAAAATTAGAAGAGTCTCCTGATAATGATGGTTTCTGAG 1907
Qy 201 AlaProCysArgMetLysValSerLeuProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1908 GCTCCCTGAGAAATGAAAGTTTCTATTTCCAACTAAAGCCTTAGAATTTGATGACATGCAA 1967
Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaLeuMetGln 240
Db 1968 ACTTTCAAGCAGAGCCTCCCGAAGGACCATCTGCTTCGACCTGCCATTGAATGCAA 2027
Qy 241 LysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGln 260

Db 2028 AAGTCTCTTCCAAATAAAGCCTTGGAAATTTGAAGAAATGAACAAACATTGAGAGCAGATCAG 2087
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 2088 ATGTTCCCTTTCAGAAATCAAAACAAAGAAAGTTCGAAGAAATTTCTGGGATTTCTGAGAGT 2147
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 2148 CTCGCTGAGACTGTGTTTCACAGAAAGGATGTGTGTGTACCCCAAGGCTACACATCAAAAGAA 2207
Qy 301 MetAspLysLeuSerGlyLysLeuGluAspSerThrSerLeuSerLysLeuLeuAspThr 320
Db 2208 ATGGATAAATAAGTGGAAATTTAGAAGATTCACCTAGCCCTATCAAAATCTTGGATACA 2267
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 2268 GTTCATTCTTGTGAAGAGCAAGGAACTTCNAAGAGATCACTGTGAACACGTCACAGGA 2327
Qy 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLeuSerGluAlaLys 360
Db 2328 AAAATGGAACAAATGAAAAAGAGTTTGTGTACTGAAAAAGAACTGTTCAGAAAGCAAAA 2387
Qy 361 GluLeuLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 2388 GAAATAAATCAGACTAGAGAACCAAAAGTTAAATGGGAAAGAGCTCTGCAGTGTG 2447
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleLeuSerTyrMetLysIleAlaCys 398
Db 2448 AGTTTCTCACACTCATGAAATGAAATTTATCTTCTACATGAAATTTGCATGT 2501
RESULT 11
ABT33218
ID ABT33218 standard; DNA; 3681 BP.
XX AC ABT33218;
DT 15-MAY-2003 (first entry)
XX Human tumour-related DNA sequence - SEQ ID No 463.
DE Human; ds; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX Homo sapiens.
XX WO200283956-A1.
XX 24-OCT-2002.
XX 15-APR-2002; 2002WO-US012378.
XX 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
PA Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Panger GR;
PI Vedvick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX Example 1; Page 287-288; 375pp; English.
PS The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for

CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present DNA sequence represents a
CC human tumour-related DNA sequence

XX
SQ Sequence 3681 BP; 1411 A; 680 C; 758 G; 826 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.: 1.23e-173 Length: 3681
Score: 2023.00 Matches: 393
Percent Similarity: 98.99% Conservative: 1
Best Local Similarity: 98.74% Mismatches: 4
Query Match: 98.59% Indels: 0
DB: 10 Gaps: 0

US-09-489-079-24 (1-398) x ABT33218 (1-3681)

Qy 1 MetGlnYsSerValProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAla 20
Db 1308 ATGCAAAAGTCTGTCCTCCCAATTAAGCCTTGGAAATGAATAATGAACAACATTGAGAGCA 1367
Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 1368 GATGAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAGCTTCTTGGGATTCT 1427
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 1428 GAGAGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTTTACCAAGGCTCRKATCAA 1487
Qy 61 LysGluLeuLeuAspLysAlaLeuGluGluSerProValLysAspGlyLeuLeu 80
Db 1488 AAGAAATAGATAAAATAAATGAATTAAGAGGCTCTCTGTTAAAGATGGTCTTCTG 1547
Qy 81 LysAlaAsnCysGlyMetLysValSerLysLeuProThrLysAlaLeuGluLeuMetAspMet 100
Db 1548 AAGGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTCATGAGCATG 1607
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaLeuGluMet 120
Db 1608 CAAACTTTCAAGCAGAGACCTCCCGAAGGCATCTCGCTCGAGCCTGCCATTTGAATG 1667
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAsp 140
Db 1668 CAAAGTCTGTCTCCAAATAAGCCTTGGAAATGAAGAAATGAACAACATTGAGAGCAGAT 1727
Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1728 GAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAGTCTTGGGATCTTGAG 1787
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1788 AGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTTTACCAAGGCTCRKATCAAAA 1847
Qy 181 GluLeuAspLysAlaLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
Db 1848 GAAATAGATAAAATAAATGAATTAAGAGAGTCTCTGATATGATGTTTCTGAG 1907
Qy 201 AlaProCysArgMetLysValSerLysLeuProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1908 GCTCCCTGCAGAAATGAAAGTTCTATTCCAACTAAAGCCTTAGAATTCATGAGCATGCAA 1967
Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaLeuGluMetGln 240
Db 1968 ACTTTCAAAGCAGAGCTCCCGAAGCCATCTCGCTTCGAGCCTGCCATTTGAATGCAA 2027
Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGln 260
Db 2028 AAGTCTGTCTCCAAATAAGCCTTGGAAATGAAGAAATGAACAACATTGAGAGCATCAG 2087
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
Db 2088 ATGTTCCCTTTCAGAAATCAAAACAAAGAGAGTTCGAGAAATTTCTTGGGATCTTGAGAGT 2147

Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
Db 2148 CTGCGTGAGACTGTTTTCACAGAGAGATGTGTGTACCCAGGCTACACATCAAAAGAA 2207
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 2208 ATGCATAAAATAAGTGGAAAATTAGAAGATTCAACTAGCCTATCAAAAATCTTGGATACA 2267
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 2268 GTTCATTCTTTGTGAAAGAGCAGAGGAACTTCAAAAAGATCACGTGTGAACACGTACAGGA 2327
Qy 341 LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 2328 AAAATGGAACAATGAAGAAAGTGTGTGTACTGNAAGAAAGAACTGTCTCAGAGCAAAA 2387
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 2388 GAAATAAAATCACAGTTAGAGAACCAAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 2447
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
Db 2448 AGTTTTCTCACACTCATGAAAATGAAAATTTATCTCTTACATGAAAATTCATGT 2501
RESULT 12
ADL93125
ID ADL93125 standard; cDNA; 3681 BP.
XX
AC ADL93125;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human breast cancer-associated polypeptide cDNA #458.
XX
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human; ss;
XX
OS Homo sapiens.
XX
PN US2003166022-A1.
XX
PD 04-SEP-2003.
XX
PF 15-APR-2002; 2002US-00124805.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
(CORI-) CORIXA CORP.
XX
XX Houghton RL, Sleath PR, Persing DH;
XX
XX WPI; 2003-874918/81.
DR P-PSDB; ADL93131.
XX
PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX
PS Example 1; SEQ ID NO 463; 294pp; English.
XX
CC The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a

CC cancer in a patient and treating a cancer in a patient. The present
 CC sequence represents cDNA encoding a human breast cancer-associated
 CC polypeptide.

SQ Sequence 3681 BP; 1411 A; 580 C; 758 G; 826 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.: 1.23e-173 Length: 3681
 Score: 2023.00 Matches: 393
 Percent Similarity: 98.99% Conservative: 1
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.59% Indels: 0
 DB: 11 Gaps: 0

US-09-489-079-24 (1-398) x ADL93125 (1-3681)

QY	1	MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla	20
DB	1308	ATGCAAAAGTCTGTCCTCCAAATTAAGCCTTGGAAATTTGAAAAATGAACAAACATTGAGACG	1367
QY	21	AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTyrAspSer	40
DB	1368	GATGAGATCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATCT	1427
QY	41	GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln	60
DB	1428	GAGAGTCTCTGTGAGACTGTTTACAGAGGATGTGTGTTTACCCCAAGCTRCRCATCAA	1487
QY	61	LysGluLeuLeuAspLysLeuGlnLysGluGluSerProValLysAspClyLeuLeu	80
DB	1488	AAAGAAATAGATAAAATTAATCGAAATTTAGAAAGGCTCTCTGTTTAAAGATGGTCTTCTG	1547
QY	81	LysAlaAsnGlyMetLysValSerLeuProLysAlaLeuGluLeuMetAspMet	100
DB	1548	AAGGTAAGTCTGCGAATGAAGTTTCTTCCAACTTAAGCCTTAGAATTTGATGACATG	1607
QY	101	GlnThrPheLeuAlaGluProGluLysProSerAlaPheGluProAlaLeuMet	120
DB	1608	CAAACTTTCAAGCAGAGCCTCCGAGAGGATCTGCTTCCGAGCTGCCATTTGAAATG	1667
QY	121	GlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAsp	140
DB	1668	CAAAAGTCTGTTCCAAATTAAGCCTTGGAAATTTGAAGAAATGAACAAACATTGAGACGAT	1727
QY	141	GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTyrAspSerGlu	160
DB	1728	GAGATCTCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAG	1787
QY	161	SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys	180
DB	1788	AGTCTCTGTGAGACTGTTTACAGAGGATGTGTGTTTACCCCAAGCTRCRCATCAAAA	1847
QY	181	GluLeuAspLysLeuGlnLysGluGluSerProAspAsnAspClyPheLeuLys	200
DB	1848	GAATATAGATAAAATTAATCGAAATTTAGAAAGTCTCTGTAATGAATGTTTCTGAAG	1907
QY	201	AlaProCysArgMetLysValSerLeuProLysAlaLeuGluLeuMetAspMetGln	220
DB	1908	GTCCTCTGCAATGAAGTTTCTATTCACATTAAGCCTTAGAATTTGATGACATGCA	1967
QY	221	ThrPheLeuAlaGluProGluLysProSerAlaPheGluProAlaLeuMetGln	240
DB	1968	ACTTTCAAGCAGAGCCTCCGAGAGGATCTGCTTCCGAGCTGCCATTTGAAATGCAA	2027
QY	241	LysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGln	260
DB	2028	AAGTCTGTTCCAAATTAAGCCTTGGAAATTTGAAGAAATGAACAAACATTGAGACGATCAG	2087
QY	261	MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTyrAspSerGluSer	280
DB	2088	ATGTTTCCCTTCAGAAATCAAAACAAAGAAAGTTGAAGAAATTTCTTGGGATTTCTGAGAT	2147
QY	281	LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu	300

DB	2148	CTCCGTGAGACTGTTTACAGAGGATGTGTGTGTACCAAGGCTACACATCAAAAGAA	2207
QY	301	MetAspLysLeuSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr	320
DB	2208	ATGATTAATTAAGTGGAAATTTAGAGATTTCACTAGCTTATCAAAATCTTGATACA	2267
QY	321	IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly	340
DB	2268	GTTTCACTTCTGTGAAGAGCAAGGAACTTCAAAAGATCATCTGTGAAACAACGTACAGGA	2327
QY	341	LysMetGluGlnMetLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys	360
DB	2328	AAATTTGGAACAAATGAAAGAAAGTTTGTGTACTGAAAAAGAACTGTCAAGACAAA	2387
QY	361	GluIleLysSerGlnLeuGluAsnGlnLysValLysTyrGluGlnGluLeuCysSerVal	380
DB	2388	GAATTAATTAATCAGATTTAGAGAACCAAAAGCTTAATTTGGAAGAGCTCTGCAAGTGTG	2447
QY	381	ArgPheLeuThrLeuMetLysMetLysIleLeuSerTyrMetLysIleAlaCys	398
DB	2448	AGGTTTCTCACACTCATGAAATGAAATTTATCTTACATGAAATTTGCATGT	2501
RESULT 13			
ADE44415			
ID	ADE44415 standard; cDNA; 3681 BP.		
XX	AC ADE44415;		
XX	XX 29-JAN-2004 (first entry)		
DT	Human cDNA associated with breast cancer #458.		
DE	human; ss; Gene; breast tumour; cancer; vaccine; T cell stimulator;		
KW	T cell expander.		
OS	Homo sapiens.		
PN	US2003104366-A1.		
XX	05-JUN-2003.		
PF	17-APR-2000; 2000US-00551621.		
PR	28-DEC-1998; 98US-00222575.		
PR	02-APR-1999; 99US-00285480.		
PR	23-JUN-1999; 99US-00339338.		
PR	02-SEP-1999; 99US-00389681.		
PR	03-NOV-1999; 99US-00433826.		
XX	(JIAN/) JIANG Y.		
PA	(DILL/) DILLON D C.		
PA	(MITC/) MITCHAM J L.		
PA	(XUJ/) XU J.		
PA	(HARL/) HARLOCKER S L.		
XX	Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;		
PI	WPI: 2004-020270/02.		
DR	P-PSDB; ADE44421.		
XX	Novel isolated polypeptide comprising immunogenic portion of breast tumor protein or its variant, useful for formulating vaccines for inhibiting cancer development in a patient.		
PS	Claim 5; SEQ ID NO 463; 217pp; English.		
CC	The invention relates to an isolated polypeptide comprising at least an immunogenic portion of a breast tumour protein. The polynucleotide, its polypeptide, its antibody, a pharmaceutical composition comprising the fusion protein or the polynucleotide encoding it, a vaccine comprising the fusion protein or the polynucleotide encoding it, an isolated T cell population comprising T cells specific for a breast tumour protein, and a		

CC method for removing tumour cells from a biological sample is useful for
 CC inhibiting the development of a cancer in a patient. The polypeptide is
 CC useful for stimulating and/or expanding T cells specific for a breast
 CC tumour protein. Stimulating and/or expanding T cells specific for a
 CC breast tumour protein is useful for inhibiting the development of a
 CC cancer in a patient. The method additionally involves the step of cloning
 CC at least one proliferated cell and then administering the cloned T cells
 CC to the patient. The present sequence represents a cDNA associated with
 CC breast cancer.

XX
 SQ Sequence 3681 BP; 1411 A; 680 C; 758 G; 826 T; 0 U; 6 Other;

Alignment Scores:

Pred. No.: 1,23e-173 Length: 3681
 Score: 2023.00 Matches: 393
 Percent Similarity: 98.99% Conservative: 1
 Best Local Similarity: 98.74% Mismatches: 4
 Query Match: 98.59% Indels: 0
 DB: 12 Gaps: 0

US-09-489-079-24 (1-398) x ADE44415 (1-3681)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 1308 ATGCAAAAGTCTGCCCAATAAAGCCTTGGAATTGAAATGAACAAACATTGAGAGCA 1367
 Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 1368 GATGAGATACTCCCATCAGAAATCAACAAAGAGACTATGAAGAAAGTTCTTGCGATTC 1427
 Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
 Db 1428 GAGAGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTATCCCAAGGCTRCRCATCAA 1487
 Qy 61 LysGluLeuLeuAspLysAlaGluGluGlySerProValLysAspGlyLeuLeu 80
 Db 1488 AAGAAATAGATAAAATAAATGGAATTTAGAGGGTCTCTGTTAAAGATGGTCTTCTG 1547
 Qy 81 LysAlaAsnCysGlyMetLysValSerLysLeuProThrLysAlaLeuGluLeuMetAspMet 100
 Db 1548 AAGCTTAACATCGGAATGAAAGTTCTTATTCCTCAACTAAAGCCTTAGAATTTGATGGACATG 1607
 Qy 101 GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaLleGluMet 120
 Db 1608 CAATCTTCAAGCAGAGACCTCCGAGAGCCATCTGCTTCGAGCCTGCCATTTGAATG 1667
 Qy 121 GlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAsp 140
 Db 1668 CAAAAGTCTGTTCCAAATAAAGCCTTGGAATTTGAAGAAATGAACAAACATTGAGAGCAGAT 1727
 Qy 141 GluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
 Db 1728 GAGATACTCTCCATCAGAAATCAACAAAGAGACTATGAAGAAAGTTCTTGCGATTTCTGAG 1787
 Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
 Db 1788 AGTCTCTGTGAGACTGTTTACAGAGAGATGTGTGTATCCCAAGGCTRCRCATCAAAA 1847
 Qy 181 GluLeuAspLysLleAsnGlyLysLeuGluGluSerProAspAsnAspGlyPheLeuLys 200
 Db 1848 GAAATAGATAAAATAAATGGAATTTAGAGAGTCTCTGATATATGATGTTCTTCTGAAG 1907
 Qy 201 AlaProCysArgMetLysValSerLysLeuProThrLysAlaLeuGluLeuMetAspMetGln 220
 Db 1908 GCTCCCTGCGAATGAAAGTTCTTATTCCTCAACTAAAGCCTTAGAATTTGATGGACATCAA 1967
 Qy 221 ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaLleGluMetGln 240
 Db 1968 ACTTTCAAAGCAGAGCCTCCGAGAGCCATCTGCTTCGAGCCTGCCATTTGAATGCAA 2027
 Qy 241 LysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGln 260
 Db 2028 AAGTCTGTTCCAAATAAAGCCTTGGAATTTGAAGAAATGAACAAACATTGAGAGCAGATCAG 2087

Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSerGluSer 280
 Db 2088 ATGTTCCCTTCAGATCAACAAACAAAGAGGTTGAAGAAAATTTCTGGGATTTCTGAGAGT 2147
 Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
 Db 2148 CTCGGTGAGACTGTTTACAGAGAGGATGTGTGTGTACCCCAAGGCTACACATCAAAAGAA 2207
 Qy 301 MetAspLysLleSerGlyLysLeuGluAspSerThrSerLeuSerLysLleLeuAspThr 320
 Db 2208 ATGCATAAAATAAGTGGAAAATTTAGAAGATTCAACTAGCCTATCAAAAATTTCTGGATACA 2267
 Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
 Db 2268 GTTCAATTCCTGTGAAAGAGCAGGGAACCTTCAAAAAGATCCTGTGAACAACTGACAGGA 2327
 Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
 Db 2328 AAAATGGAACAAATGAAGAAAGTCTTGTGTACTGAAAAAGAACTGTCTCAGAAGCAAAA 2387
 Qy 361 GluLysSerGlnLeuGluAsnGlnLysValLysValLysTrpGluGlnGluLeuCysSerVal 380
 Db 2388 GAAATAAAATTCACAGTTAGAGAACCAAAAAGTTTAAATGGGAACAGAGCTCTGCAGTGTG 2447
 Qy 381 ArgPheLeuThrLeuMetLysMetLysLleLleSerTyrMetLysLleAlaCys 398
 Db 2448 AGGTTTCTCACACTCATGAATGAATAATTAATCTCTTACATGAATAATTCATGT 2501
 RESULT 14
 ID ABT33259 standard; DNA; 1953 BP.
 XX ABT33259;
 AC ABT33259;
 XX 15-MAY-2003 (first entry)
 DT 15-MAY-2003 (first entry)
 XX Human tumour-related DNA sequence - SEQ ID No 549.
 DE Human; ds; vaccine; gene therapy; T cell stimulation; T cell expansion;
 KW tumour; breast cancer; cancer; immune response stimulation.
 XX Homo sapiens.
 OS
 XX
 XX W0200283956-A1.
 PN
 XX 24-OCT-2002.
 PD
 XX 15-APR-2002; 2002WO-US012378.
 PF
 XX 13-APR-2001; 2001US-00834759.
 PR 07-DEC-2001; 2001US-00007805.
 PR 13-FEB-2002; 2002US-00076622.
 XX (CORI-) CORIXA CORP.
 PA
 XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
 PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
 PI Vedvick TS, Mcneill PD, Durham M;
 XX WPI; 2003-103376/09.
 XX
 XX New polypeptide and polynucleotide useful for stimulating and/or
 PT expanding T cells specific for a tumor protein and treating breast
 PT cancer.
 XX
 XX Example 8; Page 330-331; 375pp; English.
 PS
 XX The invention comprises a method of stimulation and/or expanding T cells
 CC specific for a tumour protein. The invention further comprises human
 CC nucleic acids and proteins that are associated with tumours (e.g. breast
 CC cancer). The method and sequences of the invention are useful for
 CC stimulating and/or expanding T cells specific for a tumour protein,

CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present DNA sequence represents a
CC human tumour-related DNA sequence

XX Sequence 1953 BP; 745 A; 342 C; 411 G; 454 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 6.9e-173 Length: 1953
Score: 2011.00 Matches: 390
Percent Similarity: 98.99% Conserved: 4
Best Local Similarity: 97.99% Mismatches: 4
Query Match: 98.00% Indels: 0
DB: 10 Gaps: 0

US-09-489-079-24 (1-398) x ABT33259 (1-1953)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluLysGlnThrIleuArgAla 20
Db 757 ATGCAAAAGTCTGTCCTCCAAATAAAGCCTTGGAAATTTGAAAAATGAACAAACATTGAGACGA 816
Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTyrAspSer 40
Db 817 GATGAGATCTCCCATCATCAGATCCAAACAAAGGACTATGAGAAAAATCTTGGGATACT 876
Qy 41 GluSerLeuLysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 877 GAGAGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTACCCCAAGGCTGGCATCAA 936
Qy 61 LysGluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 937 AAAGAAATAGATAAAATAATGGAATAATAGAAAGGCTCTCTGGTAAANATGGTCTTCTG 996
Qy 81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLysMetAspMet 100
Db 997 AAGGCTAACTCGGGAATGAAAGTTTCTATTCACACTAAGCCCTTAGAATTTGATGACATG 1056
Qy 101 GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1057 CAAACTTTCAAGCAGAGCCTCCCGAAGGCCATCTGCCTTCGAGCCTGCCATTGAAATG 1116
Qy 121 GlnLysSerValProAsnLysAlaLeuGluLysAsnGluLysGlnThrIleuArgAlaAsp 140
Db 1117 CAAAAAGTCTGTTCCAAATAAAGCCTTGGAAATTTGAAAGAAATGAACAAACATTGAGACGACAT 1176
Qy 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTyrAspSerGlu 160
Db 1177 GAGATACCTCCCATCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGATTTCTGAG 1236
Qy 161 SerLeuLysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaThrHisGlnLys 180
Db 1237 AGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTACCCCAAGGCTGGCATCAAAAA 1296
Qy 181 GluIleAspLysIleAsnGlyLysLeuGluGlySerProAspAsnAspGlyPheLeuLys 200
Db 1297 GAAATAGATAAAATAATGGAATAATAGAAAGTCTCTCTGATAATGATGGTTTCTGAGG 1356
Qy 201 AlaProCysArgMetIleValSerIleProThrLysAlaLeuGluLysMetAspMetGln 220
Db 1357 TCTCCCTCGAGAAATGAAAGTTTCTATTCACACTAAGCCCTTAGAATTTGATGACATGCA 1416
Qy 221 ThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1417 ACTTTCAAGCAGAGCCTCCCGAAGGCCATCTGCCTTCGAGCCTGCCATTGAAATGCAA 1476
Qy 241 LysSerValProAsnLysAlaLeuGluLysAsnGluLysGlnThrIleuArgAlaAspGln 260
Db 1477 AAGTCTGTTCCAAATAAAGCCTTGGAAATTTGAAAGAAATGAACAAACATTGAGACGATCAG 1536
Qy 261 MetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTyrAspSerGluSer 280
Db 1537 ATGTTCCCTTCAGATTCAAACAAAGAACGTTGAGAAAAATCTTGGGATTTCTGAGAGT 1596
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300

Db 1597 CTCGTCGAGACTGTTTACAGAAAGGATGTGTGTGTACCCAGGCTACACATCAAAAAAGAA 1656
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 1657 ATGGATTAATAATAGTGGAAATTTAGAAGATTCACTAGCCTTCAAAAAATCTTGGATACA 1716
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1717 GTTCATTTCTGTGAAAGAGCAAGGAACTTCAAAAAAGATCACTGTGAACACGATACAGGA 1776
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLysLysLysLeuSerGluAlaLys 360
Db 1777 AAAATGGAACAAATGAAAAAGAGTTTGTGTACTGAAAAAGAACTGTGAGAAAGCAANA 1836
Qy 361 GluIleLysSerGlnLeuGluAsnGlnLysValLysTyrGluGlnGluLeuCysSerVal 380
Db 1837 GAAATAAATCAGACTTAGAGAACCAAAAGTTAAATGGGAACAAGAGCTCTGCAGTGTG 1896
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
Db 1897 AGGTTTCTCACACTCATGAAATGAAAAATTTATCTCTTACATGAAAAATTCATGT 1950
RESULT 15
ADL93211 standard; cDNA; 1953 BP.
XX ID ADL93211
XX AC ADL93211;
XX DT 20-MAY-2004 (first entry)
XX DE Human breast cancer-associated polypeptide cDNA #493.
XX KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human; ss;
XX OS Homo sapiens.
XX PN US2003166022-A1.
XX PD 04-SEP-2003.
XX PF 15-APR-2002; 2002US-00124805.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX PA (CORI-) CORIXA CORP.
XX PI Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.
XX P-PSDB; ADL93214.
XX PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX treating breast cancer.
XX PS Example 8; SEQ ID NO 549; 294pp; English.
XX CC The invention relates to an isolated breast cancer-associated
XX polypeptide. The polypeptide may be used for the diagnosis and treatment
XX of breast cancers. The methods are useful for detecting the presence of a
XX cancer in a patient and treating a cancer in a patient. The present

CC sequence represents cDNA encoding a human breast cancer-associated
CC polypeptide.

SQ Sequence 1953 BP; 745 A; 342 C; 411 G; 454 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 6.9e-173 Length: 1953
Score: 2011.00 Matches: 390
Percent Similarity: 98.99% Conservative: 4
Best Local Similarity: 97.99% Mismatches: 4
Query Match: 98.00% Indels: 0
DB: 11 Gaps: 0

US-09-489-079-24 (1-398) x ADL93211 (1-1953)

```
Qy 1 MetGlnLysSerValProAenLysAlaLeuGluLeuLysAenGluInThrLeuArgAla 20
Db 757 ATGC AAAAGTCTGTCCCAATAAAGCCCTTGGAATTGAAAAATGAACAAACATTGAGAC 816
Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 817 GATGAGATACCTCCCATCAGAAATCCAAACAAAGGACTATGAAGAAAAATCTTGGGATACT 876
Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 877 GAGAGTCTCTGTGAGACTGTTTACACAGAGGATGTGTGTTTACCCAGGCTGGCATCAA 936
Qy 61 LysGluIleAspLysIleAenGlyLysLeuGluGlySerProValLysAspGlyLeuLeu 80
Db 937 AAAGAAATAGATAAAATAAATGGAATAATTAGAAGGGTCTCTCGTAAANATGGTCTTCTG 996
Qy 81 LysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 997 AAGGCTAACTCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAAATTTAGTGGACATG 1056
Qy 101 GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 1057 CAACACTTTC AAGCAGAGCCCTCCGAGAGCCATCTCGCTTCGAGCCTGCCATTGAAATG 1116
Qy 121 GlnLysSerValProAenLysAlaLeuGluLeuLysAenGluInThrLeuArgAlaAsp 140
Db 1117 CAAAAGTCTGTTC CAAATAAAGCCTTGGAATTTGAAGAAATGAACAAACATTGAGAGCAGAT 1176
Qy 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 1177 GAGATACCTCCATCAGAAATCCAAACAAAGACTATGAAGAAAGTTCTTGGGATTCGAG 1236
Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLys 180
Db 1237 AGTCTCTGTGAGACTGTTTACACAGAGGATGTGTGTTTACCCAGGCTGGCATCAAAA 1296
Qy 181 GluIleAspLysIleAenGlyLysLeuGluGlySerProAspAenAspGlyPheLeuLys 200
Db 1297 GAAATAGATAAAATAAATGGAATAATTAGAAGAGTCTCTCTGATATGATGGTTTCTGAAG 1356
Qy 201 AlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 220
Db 1357 TCTCCCTGCGAGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAAATTTAGTGGACATCAA 1416
Qy 221 ThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMetGln 240
Db 1417 ACTTTCAAAGCAGAGCCCTCCGAGAGGCCATCTGCCCTTCGAGCCTGCCATTGAAATGCAA 1476
Qy 241 LysSerValProAenLysAlaLeuGluLeuLysAenGluInThrLeuArgAlaAspGln 260
Db 1477 AAGTCTGTTC CAAATAAAGCCTTGGAATTTGAAGAAATGAACAAACATTGAGAGCAGATCAG 1536
Qy 261 MetPheProSerGluSerLysGlnLysAenValGluGluAenSerTrpAspSerGluSer 280
Db 1537 ATGTTCCTTCAGAAATCAAAACAAAGACCTTGAAGAAATTTCTGGGATTTCTGAGAGT 1596
Qy 281 LeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGlnLysGlu 300
```

Search completed: August 1, 2005, 23:17:51

Job time : 774.055 secs

```
Db 1597 CTCCTGTGAGACTGTTTACACAGAGGATGTGTGTGTACCCCAAGGCTACACATCAAAAAGAA 1656
Qy 301 MetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeuAspThr 320
Db 1657 ATGGATAAAATAAGTCGAAAAATTAGAAGATTCAACTAGCCTATCAAAAAATCTTTGGATACA 1716
Qy 321 IleHisSerCysGluArgAlaArgGluLeuGlnLysAspHisCysGluGlnCysThrGly 340
Db 1717 GTTCATTCTTGTGAAAGACGACGAGGAACTTCAAAAAGATCACTGTGAACAAACGTACAGGA 1776
Qy 341 LysMetGluGlnMetLysLysLysPheCysValLeuLeuLysLysLysLeuSerGluAlaLys 360
Db 1777 AAAATGGAACAAATGAAAAAGAAAGTTTGTGTACTGAAAAAGAAACTGTTCAGAAAGCAAAA 1836
Qy 361 GluIleLysSerGlnLeuGluAenGlnLysValLysTrpGluGlnGluLeuCysSerVal 380
Db 1837 GAAATAAAATTCACAGTTAGAGAACCAAAAAGTTTAAATGGGAACCAAGAGCTCTGCAGTGTG 1896
Qy 381 ArgPheLeuThrLeuMetLysMetLysIleIleSerTyrMetLysIleAlaCys 398
Db 1897 AGGTTTCTCACACTCATGAAATGAAANATTAATCTCTTACATGAAANATTGCATGT 1950
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:44:18 ; Search time 15137.7 Seconds
(without alignments)
1000.786 Million cell updates/sec

Title: US-09-489-079-24
Perfect score: 2052
Sequence: 1 MQKSVNKALEKNEQLRA.....SVREFTLMKMKIISYMKIAC 398

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlP
-Q=/cgm2_1/USPTO_spool_p/US09489079/runat_01082005_141250_13309/app_query.fasta_1.1635
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1 1 10973 @runat_01082005_141250_13309 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT_DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gesi:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1015	49.5	3443	BC028407	BC028407 Homo sapi
2	971	47.3	582	BP313026	BP313026 BP313026
3	962	46.9	582	BP313436	BP313436 BP313436
4	922	44.9	581	BP314867	BP314867 BP314867
5	882	43.0	582	BP313704	BP313704 BP313704
6	865	42.2	574	BP328582	BP328582 RC5-BN019
7	830	40.4	582	BP314260	BP314260 BP314260
8	813	39.6	582	BP315089	BP315089 BP315089
9	806	39.3	490	AI951118	AI951118 wx63g05.x

LOCUS	BC028407	3443 bp	mrna	linear	HTC 18-MAY-2004
DEFINITION	Homo sapiens ankyrin repeat domain 30B, mRNA (CDNA clone IMAGE:4821910), containing frame-shift errors.				
ACCESSION	BC028407.1 GI:22382096				
VERSION	BC028407				
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3443)				
AUTHORS	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hong, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butler, M.C., Schmutz, J., Krzywicki, M.I., Skalek, U., Smalley, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 3443)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: gcgaps-rc@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbioology.org>
Contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 46 Row: 1 Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES
source Location/Qualifiers
1 . 3443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4821910"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"
ORIGIN
Alignment Scores:
Pred. No.: 2,69e-94 Length: 3443
Score: 1015.00 Matches: 208
Percent Similarity: 66.94% Conservative: 39
Best Local Similarity: 56.37% Mismatches: 83
Query Match: 49.46% Indels: 39
DB: 3 Gaps: 3
US-09-489-079-24 (1-398) x BC028407 (1-3443)
QY 1 MetGlnlySerValProAenLySAlaLeuGluLeuLySAnGluGlnThrLeuArgAla 20
Db 1700 ATCCAAAGAGCTGTTCAAATAAAGCTTTGAATTGAAGAAATGAACAAACATTGAGAGCA 1759
QY 21 AspGluIleLeuProSerGluSerLySglnLyAspTyrGluGluSerSerTtpAspSer 40
Db 1760 GCTCAGATGTCCATCAGAAATCCAAACAAAGGACGATGAGAGAAATTTCTTGGATTCT 1819
QY 41 GluSerLeuCySglnThrValSerGlnLyAspValCySLeuProLySAlaAlaHisGln 60
Db 1820 GAGAGTCCCTGTGAGACGGTTTCACAGAAGGATGTGTATTACCCAAAGCTACACATCAA 1879
QY 61 LysGluIleAspLySIlleAnGlnLyLySLeuGluGly----- 72
Db 1880 AAGAATTTCGATACCTTAAGTGGAAATTAGAACCTTACCTGTGGAAGGAAGTTCTCT 1939
QY 73 -----SerProVally 76
Db 1940 TCCAAATAAGCCTTAGAATTAAGGACAGAGAAACATTCAAGACGAGTCTCTGTATA 1999
QY 76 sApGlyLeuLeuLySAlaAnCySglnMetLySValSerIleProThrLySAlaLeuG 96

Db	2000	AGATGGTCTTCTGAAGCCTACCTGTGTGAAGGAAGTTTCTCTCTCCAAATATAAGCCTTAGA	205
Qy	96	uLeuMetAspMetGlnThrPhelysAlaGluProProGluLysProSerAlaPheGluPr	116
Db	2060	ATTAAGGACAGAGAAACACTCAAGACAGAGTCTCCTGATAATGATGGTCTTCTGTAAGCC	2119
Qy	116	oAlaIleGluMetGlnLysSerValProLenLysAlaLeuGluLeuLysAsnGluGlnTh	136
Db	2120	TACCTGTGGAAGGAAAGTTTCTCTCCAAATAAAGCTTTAGAAATTGAAGCAGAGAAAC	2179
Qy	136	rLeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAspPyrGluGluSerSe	156
Db	2180	ATTCAAGCAGCTTCAGATGTTCCATCAGAAATCCAAACAAAGATGATGAAGAAATTC	2239
Qy	156	rTPAspSerGluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAl	176
Db	2240	TTGGGATTTTGAGAGTTTCTTGAGGCTCTCTTACAGAAATGATGGGTGTTTACCRAAGC	2299
Qy	176	aThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGluSerProAspLenAs	196
Db	2300	TACACATCAAAAGAAATTCGATACCTTAAGTGAAATTTAGAAAGAGTCTCTCTGATAAGA	2359
Qy	196	pGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLe	216
Db	2360	TGTCCTTCTGAAGCCCTACCTGTGGAAGGAAGTTTCTCTCCAAATAAAGCCTTAGAANT	2419
Qy	216	uMetAspMetGlnThrPhelysAlaGluProProGluLysProSerAlaPheGluProAl	236
Db	2420	AAAGGACAGAAACACTCAAGCAGAGTCTCCTGATAAGATGGTCTTCTGAAGCCTAC	2479
Qy	236	alLeuGluMetGlnLysSerValProLenLysAlaLeuGluLeuLysAsnGluGlnThrLe	256
Db	2480	CTGTGTAAGGAAGTTTCTCTCCAAATAAAGCCTTAGAATTTAAAGCAGAGAAACATTT	2539
Qy	256	uArgAlaAspGlnMetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTr	276
Db	2540	AAAAGCAGCTCAGATGTTCCATCAGAAATCCAAACAAAGAGTGTGAAGAAATTTCTTG	2599
Qy	276	pAspSerGluSerLeuArgGluThrValSerGlnLysAspValCysValProLysAlaTh	296
Db	2600	GGATTTTGAGAGTTTCTCTGAGACTCTCTTACAGAAATGATGTGTGTTTACCRAAGCCTAC	2659
Qy	296	rHisGlnLysGluMetAspLysIleSerGlyLysLeuGluAAspSerThrSerLeuSerLy	316
Db	2660	ACATCAAAAGAAATTCGATACCTTAAGTGAAATTTAGAAATTTTCAGGCGGGCACT--	2717
Qy	316	sIleLeuAspThrIleHisSerCysGluArgAla-----	327
Db	2718	-----GTGGTTTCAGCGCTGTAAATCCAGCCCTTTGGGAGGCAGAGGCATGGCAT	2767
Qy	328	----ArgGluLeuGlnLysAspHis 334	
Db	2768	CACGAGGTCAGCAGATCGAGACCAT 2792	
RESULT 2			
BP313026			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezukui@ims.u-tokyo.ac.jp.

FEATURES

source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR01939"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 7,89e-91 Length: 582
Score: 971.00 Matches: 189
Percent Similarity: 97.93% Conservative: 0
Best Local Similarity: 97.93% Mismatches: 4
Query Match: 47.32% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-24 (1-398) x BP313026 (1-582)

QY 5 ValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAlaAspGluLeu 24
DB 1 GTTCCAAATAAGCCTTGGAAATGAAGATGACAAACATTTGAGAGCAGATGATCTC 60
QY 25 ProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLysCys 44
DB 61 CCATCAGAAATCCAAACAAAGGACTATGAGAAAGTCTTGGGATCTTGAGAGTCTCTGT 120
QY 45 GluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGlnLysGluLeuAsp 64
DB 121 GAGACTGTTTTCACAGAGGATGTGTGTTTACCAAGGCTGCGCATCAAAAAGAAATAGAT 180
QY 65 LysIleAenGlyLysLeuGluGlySerProValLysAspGlyLeuLysAlaAenCys 84
DB 181 AAAATAATGGAAATTAAGAGGCTCTCTGTTAAGATGGTCTTCTGAGGCTAACTGTC 240
QY 85 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 104
DB 241 GGAATGAAAGTTTCTATCCAACTAAGCCTTAGAATGATGACATGCAAACTTTTCAA 300
QY 105 AlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGlnLysSerVal 124
DB 301 GCAGAGCCTCCGAGAGCCATCTGCCTTCGAGCCCTGCCATTTGAAATGCAAAAGTCTGT 360
QY 125 ProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAlaAspGluLeuPro 144
DB 361 CCAATAAAGCCTTGGAAATGAAGATGAACAAACATTTGAGAGCAGATGAGATATCTCCA 420
QY 145 SerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLysCysGlu 164
DB 421 TCAGATCCAAACAAAGGACTATGAGAAAGTCTTGGGATCTTGAGAGTCTCTGTGAG 480
QY 165 ThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluLeuAspLys 184
DB 481 ACTGTTTTCACAGAGGATGTGTGTTTACCAAGGCTGCGCATCAAAAAGAAATAGATAA 540
QY 185 IleAenGlyLysLeuGluGluSerProAenAspGly 197
DB 541 ATAAATGGAAATTAAGAGGCTCTCTGTTAAGATGGT 579

RESULT 3

BP313436
LOCUS BP313436 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION cDNA clone OPR03209, mRNA sequence.
ACCESSION BP313436
VERSION BP313436.1 GI:52242411

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 582)

AUTHORS

Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE

Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions

JOURNAL

Genome Res. 14 (9), 1711-1718 (2004)

COMMENT

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yezukui@ims.u-tokyo.ac.jp.

FEATURES

source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR03209"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 5,84e-90 Length: 582
Score: 962.00 Matches: 187
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 46.88% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-24 (1-398) x BP313436 (1-582)

QY 3 LysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAlaAspGlu 22
DB 20 AAGTCTGTTCCAAATAAGCCTTGGAAATGAAGATGAACAAACATTTGAGAGCAGATGAG 79
QY 23 IleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSer 42
DB 80 ATACTCCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTCTTGGGATTTCTGAGAGT 139
QY 43 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGlnLysGlu 62
DB 140 CTCGTGAGACTGTTTTCACAGAGGATGTGTGTTTACCCAGGCTGCGCATCAAAAAGNA 199
QY 63 IleAspLysIleAenGlyLysLeuGluGlySerProValLysAspGlyLeuLeuLysAla 82
DB 200 ATAGATAAATAAATGGAATTTAGAGGCTCTCTCTGTTAAAGATGGTCTTCTGAAGGCT 259
QY 83 AenCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThr 102
DB 260 AACTCGCGAATGAAAGTTTCTATTCCAACTAAGCCTTAGAATTTGATGACATGCAAACT 319
QY 103 PheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGlnLys 122
DB 320 TTCAAAGCAGAGCCTCCCGAGAGCCATCTGCCTTCGAGCCCTGCCATTTGAAATGCAAAAG 379
QY 123 SerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuArgAlaAspGluLeu 142
DB 380 TCTGTTCCAAATAAAGCCTTGGAAATGAAGATGAACAAACATTTGAGAGCAGATGAGATA 439
QY 143 LeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeu 162
DB 440 CTCCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTCTTGGGATTTCTGAGAGTCTC 499
QY 163 CysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGluLeu 182
DB 500 TGTGAGACTGTTTTCACAGAGGATGTGTGTTTACCCAGGCTTACACATCAAAAAGAAATA 559

```

Qy 183 AspLysIleAenGlyLysLeu 189
Db 560 GATAAATAAATGGAATAATTA 580

RESULT 4
BP314867
LOCUS BP314867 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION cDNA clone OFR07276, mRNA sequence.
ACCESSION BP314867
VERSION BP314867.1 GI:52243842
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 581)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .581
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR07276"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.: 1.01e-85 Length: 581
Score: 922.00 Matches: 179
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 44.93% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-24 (1-398) x BP314867 (1-581)

Qy 1 MetGlnLysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuAtgAla 20
Db 43 ATGCAAAAGTCTGTTCACAAATAAGCCCTTGAATTTGAAGATGAACAAACATTGAGAGCA 102
Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 103 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 162

Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 163 GAGAGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTTTACCCAAAGGCTGCATCAA 222

Qy 61 LysGluIleAspLysIleAenGlyLysLeuGluClySerProValLysAspGlyLeuLeu 80
Db 223 AAAGAAATAGATAAAATAAATGGAATAATTAAGGGTCTCTGTGTAAGATGGTCTTCTG 282

Qy 81 LysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
Db 283 AAGCTAACTGCGGAATGAAAGTTCTTATTCACAACTAAGCCCTAGAAATGATGACATG 342

Qy 101 GlnThrPheLysAlaGluProProGluLysProSerAlaPheGluProAlaIleGluMet 120
Db 343 CAAACTTTCAAAGCAGAGCCCTCCGAGAGGCAATCTGCCTTCGAGCCCTGCCATTGAAATG 402

```

```

Qy 121 GlnLysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuAtgAlaAsp 140
Db 403 CAAAAGTCTGTTCACAAATAAGCCCTTGAATTTGAAGATGAACAAACATTGAGAGCAGAT 462

Qy 141 GluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db 463 GAGATACCTCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCTGAG 522

Qy 161 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGln 179
Db 523 AGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTTTACCCAAAGGCTGCATCAA 579

RESULT 5
BP313704
LOCUS BP313704 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION cDNA clone OFR03974, mRNA sequence.
ACCESSION BP313704
VERSION BP313704.1 GI:52242679
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .582
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR03974"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.: 1.5e-81 Length: 582
Score: 882.00 Matches: 170
Percent Similarity: 97.79% Conservat: 7
Best Local Similarity: 93.92% Mismatches: 4
Query Match: 42.98% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-24 (1-398) x BP313704 (1-582)

Qy 1 MetGlnLysSerValProAenLysAlaLeuGluLeuLysAenGluGlnThrLeuAtgAla 20
Db 39 ATGCAAAAGTCTGTTCACAAATAAGCCCTTGAATTTGAAGATGAACAAACATTGAGAGCA 98

Qy 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 99 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 158

Qy 41 GluSerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaAlaHisGln 60
Db 159 GAGAGTCTCTGTGAGACTGTTTACAGAAAGGATGTGTGTTTACCCAAAGGCTGCATCAA 218

Qy 61 LysGluIleAspLysIleAenGlyLysLeuGluClySerProValLysAspGlyLeuLeu 80
Db 219 AAAGAAATAGATAAAATAAATGGAATAATTAAGGGTCTCTGTGTAAGATGGTCTTCTG 278

```


81 LysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMet 100
279 AAGGCTAACTCGGAATCAAAGTATCTATTCAACTAAAGCTTACAAATTGATGACATG 338
Qy GlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMet 120
339 CAAACTTTTCATAGCACACGCTCCGAGAAAGCATCTGCCTTCGAGCGCTGCATTGAAATG 398
Db CAAACTTTTCATAGCACACGCTCCGAGAAAGCATCTGCCTTCGAGCGCTGCATTGAAATG 398
Qy GlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLysAlaAsp 140
121 GlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLysAlaAsp 140
Db CAAAGTCTGTTCCAAATAGAGCTTGGAAATGAAGATGAACAACTTCGAGGACAT 458
399 CAAAGTCTGTTCCAAATAGAGCTTGGAAATGAAGATGAACAACTTCGAGGACAT 458
Qy GluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
141 GluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGlu 160
Db GAGATACCTCCCATCACACTCCAGACAAAGGACTATGATGATGATGATGATGATGATGAT 518
459 GAGATACCTCCCATCACACTCCAGACAAAGGACTATGATGATGATGATGATGATGATGAT 518
Qy SerLeuGluGluThrValSerGlnLysAspValLysLeuProLysAlaThrHisGlnLys 180
161 SerLeuGluGluThrValSerGlnLysAspValLysLeuProLysAlaThrHisGlnLys 180
Db AGTCTCTGAGACTGTGTCTCAGAAAGGATGTGTCTTACCAAGGCTGCGCTTCAGAAA 578
519 AGTCTCTGAGACTGTGTCTCAGAAAGGATGTGTCTTACCAAGGCTGCGCTTCAGAAA 578
Qy Glu 181
181 Glu 181
Db GAG 581
579 GAG 581

RESULT 6
BF328582
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BF328582 574 bp mRNA linear EST 22-NOV-2000
RC5-BN0192-010900-025-E09 BN0192 Homo sapiens cDNA, mRNA sequence.
BF328582
BF328582.1 GI:11299317
EST.
Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?Cl=RC5&t2=RC5-BN0192-
010900-025-E09&t3=2000-09-01&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 25
High quality sequence stop: 525.
Location/Qualifiers
1..574
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0192"
/note="Organ: breast normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application

FEATURES
source

No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Alignment Scores: 8.71e-80 Length: 574
Pred. NO.: 865.00 Matches: 172
Score: 96.67% Conservative: 2
Percent Similarity: 95.56% Mismatches: 5
Best Local Similarity: 42.15% Indels: 1
Query Match: 2 Gaps: 0
DB:

US-09-489-079-24 (1-398) x BF328582 (1-574)
Qy 137 LeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSer 156
Db 37 TTGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 96
Qy 157 TrpAspSer-GluSerLeuGluThrValSerGlnLysAspValLysLeuProLysAla 176
Db 97 TGGGGGGCGTGGAGAGTCTCTGTGAGAGTCTTTCACAGAAGGATGTGTCTTACCAAGGC 156
Qy 176 aThrHisGlnLysGluLeuAspLysIleAsnGlyLysLeuGluGluSerProAspAsnAs 196
Db 157 TACGCGTCAAAAGAAATAGATAAAATAAATGGAATTTAGAAAGGATCTCTCTGATAATGA 216
Qy 196 pGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLe 216
Db 217 TGGTTTCTGAAGGCTCCCTGCAGAAATGAAATTTCTTATTCACAAAGGCTTGAATTT 276
Qy 216 uMetAspMetGlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAla 236
Db 277 GATGACATGCAAACTTTCAAAGCAGAGCCCTCCGAGAAAGCATCTGCTTCGAGCGCTGC 336
Qy 236 aIleGluMetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrIle 256
Db 337 CATTTGAAATGCAAAAGTCTGTTCAAAATAAAGCCCTGGAATTTGAAGAATGAACAACAT 396
Qy 256 uArgAlaAspGlnMetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTr 276
Db 397 GAGAGCAGATCAGATGTTTCCCTTCAGAAATCAAAAGGATTTGAAGAAATTTCTTG 456
Qy 276 pAspSerGluSerLeuArgGluThrValSerGlnLysAspValLysValProLysAlaTh 296
Db 457 TGATTCGAGAGTCTCCGTGAGAGTCTTTCACAGAAGGATGTGTGTGTACCAAGGCTAC 516
Qy 296 rHisGlnLysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSer 315
Db 517 ACATCAAAAGATATGGATAAAATAAGTGAATAATTTAGAGATTCAACTGCGCTATCA 574

RESULT 7
BF314260
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BF314260 582 bp mRNA linear EST 17-SEP-2004
BP314260 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
cDNA clone OFR05703, mRNA sequence.
BP314260
BP314260.1 GI:52243235
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasuku@ims.u-tokyo.ac.jp.

FEATURES

source
1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR05703"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 3.96e-76 Length: 582
Score: 830.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 40.45% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-24 (1-398) x BP314260 (1-582)

Qy 72 GlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePro 91
Db 102 GGGTCTCTCTTAAAGATGCTCTTGAAGGCTAACTCGGAATGAAAGTTTCTATTCCA 161
Qy 92 ThrLysAlaLeuLeuMetAspMetGlnThrPhelLysAlaGluProGluLysPro 111
Db 162 ACTAAAGCCTTAGAATTGATGGCATGCCAACTTTTCAAGCAGAGCCTCCCGAGAAGCCA 221
Qy 112 SerAlaPheGluProAlaIleGluMetGlnLysSerValProLysAlaLeuGluLeu 131
Db 222 TCTGCCCTTCAGAGCTGCATTTGAATGCAAAAGTCTGTTCCAAATAAGGCTTTGGAATTG 281
Qy 132 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAsp 151
Db 282 AAGNATGAACAAACATTTGAGAGCAGATGAGATCTCCCATCAGNATCCAAACAAAGGAC 341
Qy 152 TyrGluGluSerTrpAspSerGluSerLeuCysGluThrValSerGlnLysAspVal 171
Db 342 TATGAGAAAGTTCTTGGGATTTCTGAGAGTCTCTGTGAGACTGTTTTCACAGAAGGATGTG 401
Qy 172 CysLeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGlu 191
Db 402 TGTTTACCCAGGCTACACATCAAAAAGAAATAGATAAAATAATGGAAATTAGAAGAG 461
Qy 192 SerProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThr 211
Db 462 TCTCCTGATATGATGTTTCTGAAGGCTCCCTGCAGATGAAAGTTTCTATTCCAACT 521
Qy 212 LysAlaLeuGluLeuMetAspMetGlnThrPhelLysAlaGluProGluLysProSer 231
Db 522 AAAGCCTTAGAATTGATGGCATGCCAACTTTCAAGCAGAGCCTCCCGAGAAGCCATCT 581

RESULT 8
BP315089
LOCUS
DEFINITION
cDNA clone OPR07766, mRNA sequence.
ACCESSION
BP315089
VERSION
BP315089.1 GI:52244064
KEYWORDS
EST.
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 582)
Suzuki,Y., Yamaehita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions

JOURNAL
COMMENT

Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasuku@ims.u-tokyo.ac.jp.

FEATURES
source

1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR07766"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 2.35e-74 Length: 582
Score: 813.00 Matches: 160
Percent Similarity: 94.15% Conservative: 1
Best Local Similarity: 93.57% Mismatches: 10
Query Match: 39.62% Indels: 0
DB: 5 Gaps: 0
US-09-489-079-24 (1-398) x BP315089 (1-582)

Qy 42 SerLeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaHisGlnLys 61
Db 70 AGTCTCTGTGAGACTGTTTTCACAGAAGGATGTGTGTACCCCAAGGCTCGCATCAAAA 129
Qy 62 GluIleAspLysIleAsnGlyLysLeuGluGlySerProValLysAspGlyLeuLys 81
Db 130 GAAATAGATAAAATAAATGGAATTAGAAGGCTCTCTGTTAAAGATGCTCTCTGAAG 189
Qy 82 AlaAsnCysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGln 101
Db 190 GCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTAGAATTTGATGGACATCAA 249
Qy 102 ThrPhelLysAlaGluProGluLysProSerAlaPheGluProAlaIleGluMetGln 121
Db 250 ACTTTCAAAGCAGAGCCTCCCGAGAAGCCATCTGCTCGAGCTCGCATTTGAAATGCAA 309
Qy 122 LysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAlaAspGlu 141
Db 310 AAGTCTGTTCCAAATAAAGCCTTGGAAATGGAAGATGAACAAACATTTGAGAGCATGAG 369
Qy 142 IleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSer 161
Db 370 ATACTCCCATCATTAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGATTCTCTGACAGT 429
Qy 162 LeuCysGluThrValSerGlnLysAspValCysLeuProLysAlaThrHisGlnLysGlu 181
Db 430 CTCTGTGAGACTGTTTTCACAGAAGGATGTGTGTTTACCCCAAGGCTCGCATCAAAAAGAA 489
Qy 182 IleAspLysIleAsnGlyLysLeuGluGlySerProAspAsnAspGlyPheLeuLysAla 201
Db 490 ATAGATAAAATAAATAATGGAATTAGAAGGCTCTCTGTTTAAAGATGCTCTTCTGAAGGCT 549
Qy 202 ProCysArgMetLysValSerIleProThrLys 212
Db 550 AACTGCGGAATGAAAGTTTCTATTCCCAACTAAA 582
RESULT 9
AI951118/c
AI951118
LOCUS
DEFINITION
wx63g05.x1 NCI_CGAP_Br18 Homo sapiens cDNA clone IMAGE:2548376 3',
mRNA sequence.
ACCESSION
AI951118
VERSION
AI951118.1 GI:5743428
KEYWORDS
EST.
Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 490)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center

FEATURES
source
1..490
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2548376"
/tissue_type="four pooled high-grade tumors, including two primary tumors and two metastatic to ovary"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Br18"
/notes="Organ: breast; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies."

ORIGIN
Alignment Scores:
Pred. No.: 9,81e-74 Length: 490
Score: 806.00 Matches: 156
Percent Similarity: 97.55% Conservative: 3
Best Local Similarity: 95.71% Mismatches: 4
Query Match: 39.28% Indels: 0
DB: 1 Gaps: 0

US-09-489-079-24 (1-398) x AI951118 (1-490)

QY 179 GlnLysGluIleAspLysIleAenGlyLysLeuGluSerProAspAenAspGlyPhe 198
Db 489 CAAAAAGAAATAGATAAAATTAATGGAAATATGAGAGAGTCTCTGTAATATGATGTTTT 430
QY 199 LeuLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMetAsp 218
Db 429 CTGAAGGCTCCCTGCAGATGAAGTTTCATTCCAACTTAAGCCTTAGAATTGATGGAC 370
QY 219 MetGlnThrPheLysAlaGluProGluLysProSerAlaPheGluProAlaIleGlu 238
Db 369 ATGCAAACTTTCAAAGCAGAGCCTCCCGAGAGGCACTCGCTTCGAGCCTGCCATTGAA 310
QY 239 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 258
Db 309 ATGCAAAGTCTGTGTCAAATAAAGCCTTCGAAATTGAAGATGAACAAACATTGAGAGCA 250
QY 259 AspGlnMetPheProSerGluSerLysGlnLysAsnValGluGluAsnSerTrpAspSer 278
Db 249 GATCAGATGTTCCTTCAGAAATCAAAACAAAGAACGTTGAAGAAATTTCTGGGATTCT 190
QY 279 GluSerLeuArgGluThrValSerGlnLysAspValCysValProLysAlaThrHisGln 298
Db 189 GAGAGTCTCCGTGAGACTGTTTTCACAGAGGATGTGTGTGTACCAAGGCTACACATCAA 130
QY 299 LysGluMetAspLysIleSerGlyLysLeuGluAspSerThrSerLeuSerLysIleLeu 318
Db 129 AAGAAATAGTAATAATAGTGGAAATATGAGAGATTCAACTGACCTATCAAAATCTTG 70

Qy	319	AspThrIleHisSerCysGluIuArgAlaGluLeuGlnLysAspHisCysGluGlnCys	338
Db	69	GATACAGTTTCATCTTCTGTGAAAGAACAAAGGAACTTCAAAAAGACCCCTGTGACCCACGT	10
Qy	339	ThrGlyLys	341
Db	9	TCAGGAAA	1
RESULT 10	BP312910	Sugano cDNA library, mammary gland	OCUB-F Homo sapiens
LOCUS	BP312910	598 bp	mRNA linear
DEFINITION	CDNA clone OFR01441	598 bp	mRNA linear
ACCESSION	BP312910	598 bp	mRNA linear
VERSION	BP312910.1	598 bp	mRNA linear
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 598)		
JOURNAL	Suzuki.Y., Yamashita.R., Shiota,M., Sakakibara.Y., Chiba,J.,		
COMMENT	Mizushima-Sugano,J., Nakai.K. and Sugano,S.		
FEATURES	Sequence comparison of human and mouse genes reveals a homologous		
source	block structure in the promoter regions		
	Genome Res. 14 (9), 1711-1718 (2004)		
	Contact: Yutaka Suzuki		
	Institute of Medical Science, University of Tokyo		
	4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan		
	Email: yusuzuki@img.u-tokyo.ac.jp.		
	Location/Qualifiers		
	1..598		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="OFR01441"		
	/tissue_type="mammary gland"		
	/cell_line="OCUB-F"		
	/clone_lib="Sugano cDNA library, mammary gland		
	/note="mammary gland tumor"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,038-71	Length:	598
Score:	785.00	Matches:	157
Percent Similarity:	89.39%	Conservative:	3
Best Local Similarity:	87.71%	Mismatches:	11
Query Match:	38.26%	Indels:	8
DB:	5	Gaps:	1
US-09-489-079-24 (1-398) x BP312910 (1-598)			
Qy	53	CysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAenGlyLysLeuGluGly	72
Db	84	TGCTTCATGAATCTAGTATTTTCAAAATTTTCTTA	119
Qy	73	SerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThr	92
Db	120	TCCTCTGTTAAAGATGGTCTCTTCGAAGGCTTAATGCGGAATGAAGATTTCTATTCCAAC	179
Qy	93	LysAlaLeuGluLeuMetAspMetGlnThrPhelLysAlaGluProGluLysProSer	112
Db	180	AAAGCCTTAGAATTGATGGNCATGCAAACTTTCAAGCAGAGCCTCCGAGAGCCATCT	239
Qy	113	AlaPheGluProAlaIleGluMetGlnLysSerValProAenLysAlaLeuGluLeuLys	132
Db	240	GCCTTCGAGCCTGCCATTGAAATGCAAAAGTCTGTTCCAAATTAAGCTTTGGAATTGAAG	299
Qy	133	AsnGluGlnThrLeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAspTyr	152
Db	300	AATGAACAAACATTGAGAGCAGATGAGTACTCCCATCAGAAATCAAAACAAAGACTAT	359

```

Qy 153 GluGluSerSerTtpAspSerGluSerLeuCysGluThrValSerGlnLysAspValCys 172
Db 360 GAAGAAAGTTCTTGGATTCTGAGACTCTGTGAGACTGTTTACAGAGGATGTGTCT 419

Qy 173 LeuProLysAlaThrHisGlnLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 192
Db 420 TTACCCAAAGGCTGCGCATCAAAAAGAAATAGATAAAATAAATGAAAGGTCT 479

Qy 193 ProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLys 212
Db 480 CCTGTGTTAAAGATGCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAA 539

Qy 213 AlaLeuGluLeuMetAspMetGlnThrPhelysAlaGluProGluLysProSer 231
Db 540 GCCTTAGAATTGATGGACATCAAACTTCAAGCAGAGCCTCCGAGAGCCATCT 596

RESULT 11
BP315806 582 bp mRNA linear EST 17-SEP-2004
LOCUS BP315806 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION BP315806.1 GI:52244781
ACCESSION BP315806
VERSION BP315806
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 582)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..582
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR09418"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores: 5.09e-71 Length: 582
Pred. No.: 781.00 Matches: 153
Score: 95.62% Conservative: 0
Percent Similarity: 95.62% Mismatches: 7
Best Local Similarity: 38.06% Indels: 0
Query Match: 5 Gaps: 0
DB:

US-09-489-079-24 (1-398) x BP315806 (1-582)

Qy 72 GlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePro 91
Db 102 GGGTCTCTGTTAAAGATGCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCA 161

Qy 92 ThrLysAlaLeuGluLeuMetAspMetGlnThrPhelysAlaGluProGluLysPro 111
Db 162 ACTAAAGCCTTAGAATTGATGGACATCAAACTTTCAAAGCAGAGCCTCCGAGAGCCCA 221

Qy 112 SerAlaPheGluProAlaIleGluMetGlnLysSerValProAsnLysAlaLeuGluLeu 131
Db 222 TCTGCCTTCGAGCCTGCCTTGAATGAAAGTCTGTTCCAAATAAAGCCTTGGAAATTG 281

```

```

Qy 132 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAsp 151
Db 282 AAGATGAAACAACATTTGAGAGCAGATGAGATCTCCATCCAGATTCACAACAAGAGAC 341

Qy 152 TyrGluGluSerSerTtpAspSerGluSerLeuCysGluThrValSerGlnLysAspVal 171
Db 342 TATGAAGAAAGTTCTTGGATTCTGAGAGTCTCTGTGAGACTGTTTACAGAGGATGTG 401

Qy 172 CysLeuProLysAlaThrHisGlnLysGluLeuLeuLeuLeuLeuLeuLeuLeu 191
Db 402 TGTTTACCCAAAGGCTGCGCATCAAAAAGAAATAGATAAAATAAATGAAAGAGGG 461

Qy 192 SerProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThr 211
Db 462 TCTCTGTTAAAGATGCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACT 521

Qy 212 LysAlaLeuGluLeuMetAspMetGlnThrPhelysAlaGluProGluLysProSer 231
Db 522 AAAGCCTTAGAATTGATGGACATGCAAACTTTCAAAGCAGAGCCTCCGAGAGCCATCT 581

RESULT 12
BP314652 584 bp mRNA linear EST 17-SEP-2004
LOCUS BP314652 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION BP314652
ACCESSION BP314652
VERSION BP314652
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 584)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..584
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR06752"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores: 1.63e-68 Length: 584
Pred. No.: 757.00 Matches: 152
Score: 89.08% Conservative: 3
Percent Similarity: 87.36% Mismatches: 11
Best Local Similarity: 36.89% Indels: 8
Query Match: 5 Gaps: 1
DB:

US-09-489-079-24 (1-398) x BP314652 (1-584)

Qy 53 CysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGly 72
Db 83 TGCTTGATGAATCTAGTTTTTCAAAATTTTCTTA-----GGG 118

Qy 73 SerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThr 92
Db 119 TCTCCTGTTAAAGATGCTTCTTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACT 178

```

Qy 93 LysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProGluLysProSer 112
 Db 179 AAGGCTTAGAATTCATGAGCATGCAACTTTCAAAGCAGAGGCTCCCGAGAGCCATCT 238
 Qy 113 AlaPheGluProAlaIleGluMetGlnLysSerValProAsnLysAlaLeuGluLys 132
 Db 239 GCCTTCGAGCGCTGCCATGCAAAATGCAAAAGTCTGTTCCTCAAAATGAAGCCCTTGGAAATGAAG 298
 Qy 133 AsnGluGlnThrLeuArgAlaAspGluLysLeuProSerGluSerLysGlnLysAspTyr 152
 Db 299 AATGAACAACATTCAGAGCAGATGAGATCTCCCATCAGATTCCTCAAAATGAAGACATAT 358
 Qy 153 GluGluSerSerTTPAspSerGluSerLysGluThrValSerGlnLysAspValCys 172
 Db 359 GAAGAAGTCTTGGGATCTGAGAGTCTCTGTGAGAGTGTTCACAGAGGATGTGTGT 418
 Qy 173 LeuProLysAlaThrHisGlnLysGluLysLeuAspLysIleAsnGlyLysLeuGluSer 192
 Db 419 TTACCAAGGCTGCGCATCAAAAGCAATAGATAAAATGAATGGAATTAAGAGGCTCT 478
 Qy 193 ProAspAsnAspGlyPheLeuLysAlaProCysArgMetLysValSerIleProThrLys 212
 Db 479 CCTGTAAAGATGGTCTTCTGAAGGCTAACTCGGAATGAAGATGAAAGTCTTATTCCTCAACTAAA 538
 Qy 213 AlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluPro 226
 Db 539 GCCTTAGAATTCATGAGCATGCAAACTTTCAAAGCAGAGCCT 580

RESULT 13
 BP313822
 LOCUS
 DEFINITION BP313822 583 bp mRNA linear EST 17-SEP-2004
 CDNA clone OFR04566, mRNA sequence.

ACCESSION BP313822
 VERSION BP313822.1 GI:52242797
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
 source
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OFR04566"
 /tissue_type="mammary gland"
 /cell_line="OCUB-F"
 /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 4,93e-64 Length: 583
 Score: 714.00 Matches: 144
 Percent Similarity: 85.38% Conservative: 2
 Best Local Similarity: 84.21% Mismatches: 11
 Query Match: 34.80% Indels: 14
 DB: 5 Gaps: 1

US-09-489-079-24 (1-398) x BP313822 (1-583)

Qy 72 GlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePro 91
 Db 71 GGGTCTCTCTGTTAAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTCTTCTATTCCA 130
 Qy 92 ThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProGluLysPro 111
 Db 131 ACTAAAGCCTTAGAATTCATGAGCATGCAAACTTTCAAAGCAGAGCCTCCCGAGAGCCA 190
 Qy 112 SerAlaPheGluProAlaIleGluMetGlnLysSerValProAsnLysAlaLeuGluLeu 131
 Db 191 TCTGCTTCGAGCCTGCCATTCGAATGCAAAAGTCTGTTCCTCAATGAAGCCTTGGAAATG 250
 Qy 132 LysAsnGluGlnThrLeuArgAlaAspGluLysLeuProSerGluSerLysGlnLysAsp 151
 Db 251 AAGAATGAACAACAGTTCGAGAGCAGATGAGATCTCCCATCAGAATCCCAACAAAGGAC 310
 Qy 152 TyrGluGluSerSerTTPAspSerGluSerLysValSerGlnLysAspVal 171
 Db 311 TATGAAGAAGTCTTCTGGGATTCGAGAGTCTCTGTGAGAGTGTTCCTCAAGAGATGTG 370
 Qy 172 CysLeuProLysAlaThrHisGlnLysGluLysLeuAspLysIleAsnGlyLysLeuGlu 190
 Db 371 TGTTCACCAAGGCTGCCATCAAAAGCAATAGATAAAATGAATGGAATTAAGAGGT 430
 Qy 191 -----GluSerProAspAsnAspGly 197
 Db 431 ACATATGCTGCTGAATTTAGAACATTTCTCTGCCATGATAAGGTCTCTCTGTTAAAGATGGT 490
 Qy 198 PheLeuLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMet 217
 Db 491 CTTCTGATAGCTAACTGCGGAATGAAAGTCTTCTATTCCTCAACTACTGCTTAAACTGATG 550
 Qy 218 AspMetGlnThrPheLysAlaGluProGlu 228
 Db 551 GACTTGACAGATGTCATAGCAGAGCCTCCCGAG 583

RESULT 14

BP313235

LOCUS

DEFINITION BP313235 582 bp mRNA linear EST 17-SEP-2004

CDNA clone OFR02569, mRNA sequence.

ACCESSION BP313235

VERSION BP313235.1 GI:52242210

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES

source

1..582

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="OFR02569"

/tissue_type="mammary gland"

/cell_line="OCUB-F"

/clone_lib="Sugano cDNA library, mammary gland OCUB-F"

/note="mammary gland tumor"

ORIGIN

Alignment Scores:

Pred. No.: 1.99e-61 Length: 582
Score: 689.00 Matches: 138
Percent Similarity: 86.96% Conservative: 2
Best Local Similarity: 85.71% Mismatches: 7
Query Match: 33.58% Indels: 14
DB: 5 Gaps: 1

US-09-489-079-24 (1-398) x BP313235 (1-582)

Qy 72 GlySerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIlePro 91
Db 100 GGGCTCTCTGTTAAAGATGGTCTCTGAAGGCTAACTGCGGAATGAAGTTCTATTCCA 159
Qy 92 ThrLysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysPro 111
Db 160 ACTAAGCCTTAGAATTGATGCACATGCAAACTTTCAAAGCAGAGCTCCCGAGAAGCCA 219
Qy 112 SerAlaPheGluProAlaIleGluMetGlnLysSerValProAsnLysAlaLeuGluLeu 131
Db 220 TCTGCTTCGAGCGCTGCCATTGAAATGCAAAAGTCTGTTCCAATAAAGCCCTTGGAAATTG 279
Qy 132 LysAsnGluGlnThrLeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAsp 151
Db 280 AAGAAATGAACAAACATTGAGAGCAGATGAGATCTCCCATCAGAAATCAAAACAAAGGAC 339
Qy 152 TyrGluGluSerSerTrpAspSerGluSerLysGluThrValSerGlnLysAspVal 171
Db 340 TATGAAGAAAGTCTCTGGGATTCAGAGTCTCTGTGAGACTGTTTCACAGAAGGATGTG 399
Qy 172 CysLeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGlu--- 190
Db 400 TGTTTACCAGGCTACACATCAAAAGAAATAGATAAAATGAAATGAAAAATTAGAAGGT 459
Qy 191 -----GluSerProAspAsnAspGly 197
Db 460 AGATATGCTGCTGAATTAGAACATTTCTCTCAATGATAGAGACTCTCTTAAAGATGGT 519
Qy 198 PheLeuLysAlaProCysArgMetLysValSerIleProThrLysAlaLeuGluLeuMet 217
Db 520 CTTCTGAAGGCTAACTGCGAGTGAAGTTCTATTCCAACTAAGCCTTAGAATTGATG 579
Qy 218 Asp 218
Db 580 GAC 582

RESULT 15
BP312890 583 bp mRNA linear EST 17-SEP-2004
LOCUS BP312890 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION cDNA clone OPR01356, mRNA sequence.
ACCESSION BP312890
VERSION BP312890.1 GI:52241865
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 583)
Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="OPR01356"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:

Pred. No.: 3.96e-54 Length: 583
Score: 619.00 Matches: 124
Percent Similarity: 89.44% Conservative: 3
Best Local Similarity: 87.32% Mismatches: 7
Query Match: 30.17% Indels: 8
DB: 5 Gaps: 1

US-09-489-079-24 (1-398) x BP312890 (1-583)

Qy 53 CysLeuProLysAlaAlaHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGly 72
Db 83 TGCTTGATGAATCTAGTTTTCAAAATTTCTTA-----GGG 118
Qy 73 SerProValLysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThr 92
Db 119 TCTCCTGTTAAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAGTTCTTATTTCCAAC 178
Qy 93 LysAlaLeuGluLeuMetAspMetGlnThrPheLysAlaGluProProGluLysProSer 112
Db 179 AAGCCTTAGAATTGATGGACATGCCAACTTTCAAAGCAGAGCTCCCGAGAAGCCATCT 238
Qy 113 AlaPheGluProAlaIleGluMetGlnLysSerValProAsnLysAlaLeuGluLeuLys 132
Db 239 GCCTTCGAGCGCTGCCATTGAAATGCAAAAGTCTGTTCCAAATAAAGCCTTGGAAATTGAAG 298
Qy 133 AsnGluGlnThrLeuArgAlaAspGluIleLeuProSerGluSerLysGlnLysAspTyr 152
Db 299 AATGAACAAACATTGAGAGCAGATGAGATACTCCCATCAGAAATCAAAACAAAGGACTAT 358
Qy 153 GluGluSerSerTrpAspSerGluSerLeuCysGluThrValSerGlnLysAspValCys 172
Db 359 GAAGAAAGTTCTTGGGATTCAGAGTCTCTGTGAGACTGTTTCACAGAAGGATGTGTGT 418
Qy 173 LeuProLysAlaThrHisGlnLysGluIleAspLysIleAsnGlyLysLeuGluGluSer 192
Db 419 TTACCCAAAGGCTACACATCAAAAGAAATAGATAAAATAAATGGAATAATTAGAGCCATT 478
Qy 193 ProAsp 194
Db 479 ACGGAT 484

Search completed: August 2, 2005, 11:45:51

Job time : 15149.7 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:04:01 ; Search time 141.469 Seconds
(without alignment)
1088.086 Million cell updates/sec

Title: US-09-489-079-24

Perfect score: 2052

Sequence: 1 MQKSPVKNALKEKNEQTLRA.....SVRFLLTMKMKIISYMKIAC 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04.*
1: Genesep1980s.*
2: Genesep1990s.*
3: Genesep2000s.*
4: Genesep2001s.*
5: Genesep2002s.*
6: Genesep2003as.*
7: Genesep2003bs.*
8: Genesep2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2052	100.0	398	3 AAB07638	Aab07638 Amino aci
2	2023	98.6	650	4 AAB50263	Aab50263 Human bre
3	2023	98.6	650	4 AAG65983	Aag65983 B726P spl
4	2023	98.6	650	4 AAU33346	Aau33346 Human bre
5	2023	98.6	650	5 ABG78913	Abg78913 Human bre
6	2023	98.6	650	6 ABJ37736	Abj37736 Human tum
7	2023	98.6	650	7 ADL93131	Adl93131 Human bre
8	2023	98.6	650	8 ADE44421	Ade44421 Human bre
9	2023	98.6	743	4 AAU33358	Aau33358 Human bre
10	2023	98.6	743	5 ABG78925	Abg78925 Human bre
11	2023	98.6	743	6 ABJ37748	Abj37748 Human tum
12	2023	98.6	743	7 ADL93156	Adl93156 Human bre
13	2018	98.3	661	6 ABJ37782	Abj37782 Human tum
14	2018	98.3	661	7 ADL93214	Adl93214 Human bre
15	1940.5	94.6	1002	4 AAU33351	Aau33351 Human bre
16	1940.5	94.6	1002	5 ABG78918	Abg78918 Human bre
17	1940.5	94.6	1002	6 ABJ37741	Abj37741 Human tum
18	1940.5	94.6	1002	7 ADL93137	Adl93137 Human bre
19	1940.5	94.6	1002	8 ADE44427	Ade44427 Human bre
20	1940.5	94.6	1095	4 AAU33357	Aau33357 Human bre
21	1940.5	94.6	1095	5 ABG78924	Abg78924 Human bre
22	1940.5	94.6	1095	6 ABJ37747	Abj37747 Human tum
23	1940.5	94.6	1095	7 ADL93155	Adl93155 Human bre
24	1938.5	94.5	1341	4 AAB84702	Aab84702 Amino aci
25	1938.5	94.5	1341	5 ABJ05537	Abj05537 Breast ca

26	1938.5	94.5	1341	6 ABJ37784	Abj37784 Human tum
27	1938.5	94.5	1341	6 ABR47548	Ab47548 Breast ca
28	1938.5	94.5	1341	7 ADL93227	Adl93227 Human bre
29	1938.5	94.5	1349	6 ABJ37788	Abj37788 Human tum
30	1938.5	94.5	1349	7 ADL93235	Adl93235 Human bre
31	1935.5	94.3	1013	6 ABJ37783	Abj37783 Human tum
32	1935.5	94.3	1013	7 ADL93215	Adl93215 Human bre
33	1636.5	79.8	445	4 AAB50249	Aab50249 Human bre
34	1636.5	79.8	445	4 AAG65987	Aag65987 B726P spl
35	1636.5	79.8	445	4 AAU33350	Aau33350 Human bre
36	1636.5	79.8	445	5 ABG78917	Abg78917 Human bre
37	1636.5	79.8	445	6 ABJ37740	Abj37740 Human tum
38	1636.5	79.8	445	7 ADL93135	Adl93135 Human bre
39	1636.5	79.8	445	8 ADE44425	Ade44425 Human bre
40	1510.5	73.6	512	4 AAB84701	Aab84701 Amino aci
41	1408.5	68.6	466	4 AAB50248	Aab50248 Human bre
42	1408.5	68.6	466	4 AAG65986	Aag65986 B726P spl
43	1408.5	68.6	466	4 AAU33349	Aau33349 Human bre
44	1408.5	68.6	466	5 ABG78916	Abg78916 Human bre
45	1408.5	68.6	466	6 ABJ37739	Abj37739 Human tum

ALIGNMENTS

RESULT 1

AAB07638
ID AAB07638 standard; protein; 398 AA.

AC AAB07638;

XX 07-NOV-2000 (first entry)

DE Amino acid sequence of BS322 polypeptide.

KW BS322; breast tissue marker; breast disease; breast cancer.

XX Homo sapiens.

XX WO200043420-A1.

PD 27-JUL-2000.

XX 21-JAN-2000; 2000WO-US001452.

XX 21-JAN-1999; 99US-00234716.

XX (ABBO) ABBOTT LAB.

PA Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;

XX WPI; 2000-499217/44.

DR N-PSDB; AAA59015.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the detection and diagnosis of breast disease especially breast cancer.

PS Claim 23; Page 122-123; 126pp; English.

XX The present sequence represents a human BS322 polypeptide. BS322 is a breast tissue marker. The BS322 polynucleotides and polypeptides are used to detect and diagnose breast disease, e.g. breast cancer. The BS322 polynucleotides are useful as a source of probes and primers, and the BS322 polypeptides are useful as antigens

XX Sequence 398 AA;

Query Match 100.0%; Score 2052; DB 3; Length 398;

Best Local Similarity 100.0%; Pred. No. 1.9e-152;

Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 2

KASOPI 2
 AAB50263

AA50263
ID AAB50263 standard; protein; 650 AA.

XX
CONTACT:

AC AAB50263:

AC XX
FEB 20 1963

DT 13-MAR-2001 (first entry)

XX
XX
XX

TS-PMK-ZOVI (TITOC ENERGY)

Human breast

5
 5
 5
 5
 5
 4
 2
 77511597
 55 XX
 55 XX

Human: breast

[illegible]

Homo sapiens.

XX
03
ношо варгелъ.XX
PN
WO200060076-A2

FN
XX
WUZ00060076-AZ.

12-OCT-2000
XX
BX

FD-302 (Rev. 12-1-77)

XX
RE 1E-FEB-2000. 2000W0-

PF 15-FEB-2000; 2000WO-
yy

[illegible]

PR 02-APR-1999; 99US-

PR 23-JUN-1999; 99US-

PR 02-SEP-1999; 99US-

PR 03-NOV-1999; 99US-

XX

PA (CORI-) CORIXA CORP.

Query Match	98.6%;	Score 2023;	DB 4;	Length 650;
Best Local Similarity	98.7%;	Pred.No. 7.1e-150;		
Matches 393;	Conservative 1;	Mismatches 4;	Indels 0;	Gaps 0;

QY	1	MOKSVPNKALELKNEQTLRADEILPSESKQDYESSWDSSESLCETVSKQDVCLPKAAHQ	60
Db	253	MOKSVPNKALELKNEQTLRADEILPSESKQDYESSWDSSESLCETVSKQDVCLPKAXHQ	312
QY	61	KEIDKINGKLEGGSPVDGILLKANGCMKVSIPTKALELMDMTFKAEPPEKPSAFEPAIEM	120
Db	313	KEIDKINGKLEGGSPVDGILLKANGCMKVSIPTKALELMDMTFKAEPPEKPSAFEPAIEM	372
QY	121	QKSVPNKALELKNEQTLRADEILPSESKQDYESSWDSSESLCETVSKQDVCLPKATHQK	180
Db	373	QKSVPNKALELKNEQTLRADEILPSESKQDYESSWDSSESLCETVSKQDVCLPKAXHQK	432
QY	181	EIDKINGKLBESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPEKPSAFEPAIEMQ	240
Db	433	EIDKINGKLBESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPEKPSAFEPAIEMQ	492
QY	241	KSVPNKALELKNEQTLRADQMFPSSEKQKVNSESWDSSESLRETIVSKQVCVPKATHQKE	300
Db	493	KSVPNKALELKNEQTLRADQMFPSSEKQKVNSESWDSSESLRETIVSKQVCVPKATHQKE	552
QY	301	MDKISGKLBESTSLSKILDTHTHCERARELQKHCEQCTGKMEOMKKKFCVLKKKLEAK	360
Db	553	MDKISGKLBESTSLSKILDTVHSCERARELQKHCEQRTGKMEOMKKKFCVLKKKLEAK	612
QY	361	EIKSOLENQVKWEQLCSVRFLTMKMKIISYMKIAC	398
Db	613	EIKSOLENQVKWEQLCSVRFLTMKMKIISYMKIAC	650

RESULT 3

AAG65983

ID AAG65983 standard; protein; 650 AA.

XX AC AAG65983;

XX DT 11-FEB-2002 (first entry)

XX DE B726P splice variant sequence.

XX KW Genetic subtraction; DNA microarray analysis; polymerase chain reaction; cancer; B726P.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 310 /note= "any amino acid"

FT Misc-difference 429 /note= "any amino acid"

FT Misc-difference 522 /note= "any amino acid"

XX WO200175171-A2.

XX PD 11-OCT-2001.

XX PF 02-APR-2001; 2001WO-US010631.

XX PR 03-APR-2000; 2000US-0194241P.

XX PR 20-JUL-2000; 2000US-0219862P.

XX PR 27-JUL-2000; 2000US-0221300P.

XX PR 18-DEC-2000; 2000US-0256592P.

XX PA (CORI-) CORIXA CORP.

XX PI Houghton RL, Dillon DC, Molesch DA, Xu J, Zehentner B, Persing DH;

XX WPI; 2001-626449/72.

XX DR N-PSDB; AAI67218.


```

XX Identifying tissue (tumor)-specific polynucleotides overexpressed in
PT tissue of interest as compared to control tissue, for detecting cancer
PT cells in patient, comprises DNA microarray analysis or quantitative
PT polymerase chain reaction.
XX
XX Example; Page 108-111; 127pp; English.
XX
XX The invention relates to identifying tissue-specific polynucleotides (P)
XX that involves performing a genetic subtraction to identify pool of (P)
XX from tissue of interest (TI), performing DNA microarray analysis to
XX identify first subset of polynucleotides (SP1) at least 2-fold over
XX expressed in TI, and performing quantitative polymerase chain reaction
XX (PCR) analysis on SP1 to identify second subset of (P). The method is
XX useful for determining the presence or absence of a cancer cell in a
XX patient, monitoring the progression of cancer in a patient using a
XX biological sample such as blood, serum, lymph nodes, bone marrow, sputum,
XX urine or a tumour biopsy sample. The methods are useful for determining
XX the presence or absence of or monitoring progression of prostate, breast,
XX colon, ovarian, lung, head and neck, lymphoma, leukemia, melanoma, liver,
XX gastric, kidney, bladder, pancreatic or endometrial cancer. Sequences
XX AAG5983-987 represent determined splice variants of B726P
XX
XX Sequence 650 AA;
XX
XX Query Match          98.6%; Score 2023; DB 4; Length 650;
XX Best Local Similarity 98.7%; Pred. No. 7.1e-150;
XX Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 MOKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKAAHQ 60
XX DB 253 MOKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKAAHQ 312
XX
XX QY 61 KEIDKINGKLESGSPVKDGLLKANCQKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 120
XX DB 313 KEIDKINGKLESGSPVKDGLLKANCQKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 372
XX
XX QY 121 OKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKATHQK 180
XX DB 373 OKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKAXHQK 432
XX
XX QY 181 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 240
XX DB 433 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 492
XX
XX QY 241 KSVPNKALELKNQOTLRADQMPFPPSESQKQNVENSWSDESLSRETYSQKDVCPKATHQKE 300
XX DB 493 KSVPNKALELKNQOTLRADQMPFPPSESQKQNVENSWSDESLSRETYSQKDVCPKATHQKE 552
XX
XX QY 301 MDKISGKLEDSTSLSKILDTHTHSCERARELOKDHCEQCTGKQKMEQKKFCVLKXKXSEAK 360
XX
XX RESULT 4
XX AAU33346
XX ID AAU33346 standard; protein; 650 AA.
XX
XX AC AAU33346;
XX
XX XX 18-DEC-2001 (first entry)
XX
XX DE Human breast cancer protein encoded by cDNA B726P-spliced_seq_B726P.
XX
XX KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX
XX KW gene therapy; immunogen.
XX
XX OS Homo sapiens.
XX

```

```

PN WO200179286-A2.
XX
XX PD 25-OCT-2001.
XX
XX XX 12-APR-2001; 2001WO-US012164.
XX
XX PR 17-APR-2001; 2000US-00551621.
XX
XX PR 08-JUN-2000; 2000US-00590751.
XX
XX PR 22-JUN-2000; 2000US-00604287.
XX
XX PR 20-JUL-2000; 2000US-00620405.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
XX
XX N-PSDB; AAS47405.
XX
XX Breast Tumor Proteins and nucleic acids useful for the prevention,
XX diagnosis and treatment of breast cancer.
XX
XX Claim 3; Page 275-276; 297pp; English.
XX
XX The invention relates to isolated breast tumour proteins and nucleic
XX acids that encode them, including immunogenic fragments of the proteins.
XX Also included are expression vectors expressing the proteins, transformed
XX cells and antibodies raised against the proteins or an antigen presenting
XX cell expressing the protein. The proteins and nucleic acids may be used
XX in the prevention, diagnosis and treatment of diseases associated with
XX inappropriate breast tumour protein expression, i.e. breast tumours and
XX breast cancer e.g. by gene therapy. The nucleic acids and their
XX complements may also be used as DNA probes in diagnostic assays to detect
XX and quantitate the presence of similar nucleic acids in samples, and
XX therefore which patients may be in need of restorative therapy. The
XX proteins, nucleic acids and antibodies may be used in assays to identify
XX modulators (e.g. antagonists) of breast tumour protein expression and
XX regulate expression and activity. The antibodies may also be used as
XX diagnostic agents for detecting the presence of the proteins in samples
XX (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
XX purification diagnostic techniques. The present sequence is a breast
XX tumour protein encoded by a cDNA from a breast tumour cDNA library
XX isolated by subtractive hybridisation against a normal breast cDNA
XX library
XX
XX Sequence 650 AA;
XX
XX Query Match          98.6%; Score 2023; DB 4; Length 650;
XX Best Local Similarity 98.7%; Pred. No. 7.1e-150;
XX Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 1 MOKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKAAHQ 60
XX DB 253 MOKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKAAHQ 312
XX
XX QY 61 KEIDKINGKLESGSPVKDGLLKANCQKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 120
XX DB 313 KEIDKINGKLESGSPVKDGLLKANCQKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 372
XX
XX QY 121 OKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKATHQK 180
XX DB 373 OKSVPNKALELKNQOTLRADAILPSESQKQDYESSWDSLSCTVTSQKDVCLPKAXHQK 432
XX
XX QY 181 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 240
XX DB 433 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTQFKAEPPKPSAFEPATEM 492
XX
XX QY 241 KSVPNKALELKNQOTLRADQMPFPPSESQKQNVENSWSDESLSRETYSQKDVCPKATHQKE 300
XX DB 493 KSVPNKALELKNQOTLRADQMPFPPSESQKQNVENSWSDESLSRETYSQKDVCPKATHQKE 552
XX
XX QY 301 MDKISGKLEDSTSLSKILDTHTHSCERARELOKDHCEQCTGKQKMEQKKFCVLKXKXSEAK 360
XX

```

Db 553 MDKISGKLEDSTSLKILDTVHSCERARELQKHCEQRTGMEQMKKFCVLKKLSEAK 612

Qy 361 EIKSQLENOKVKEQELCSVRFLTLMMKLIISYMKIAC 398

Db 613 EIKSQLENOKVKEQELCSVRFLTLMMKLIISYMKIAC 650

RESULT 5

ABG78913

ID ABG78913 standard; protein; 650 AA.

AC ABG78913;

XX

XX 15-NOV-2002 (first entry)

DT

XX Human breast tumour polypeptide #5.

DE

XX Human; breast tumour protein; breast cancer; cytostatic; vaccine.

KW

XX Homo sapiens.

OS

XX US2002085998-A1.

PN

XX 04-JUL-2002.

PD

XX

XX 13-APR-2001; 2001US-00834759.

PF

XX 28-DEC-1998; 98US-00222575.

PR

XX 02-APR-1999; 99US-00285480.

PR

XX 23-JUN-1999; 99US-00393338.

PR

XX 02-SEP-1999; 99US-00389681.

PR

XX 03-NOV-1999; 99US-00433826.

PR

XX 17-APR-2000; 2000US-00551821.

PR

XX 08-JUN-2000; 2000US-00590751.

PR

XX 22-JUN-2000; 2000US-00604287.

PR

XX 20-JUL-2000; 2000US-00620405.

XX

XX (CORI-) CORIXA CORP.

PA

XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;

PI Henderson RA;

PI

XX WPI; 2002-635657/68.

DR

XX N-PSDB; ABS64006.

DR

XX

XX Novel breast cancer polynucleotides and polypeptides encoded by the

PT polynucleotides, useful for detecting the presence of breast cancer in a

PT patient, and in pharmaceutical compositions, for treating breast cancer.

XX

XX Claim 2; Page 201-203; 247pp; English.

PS

XX The invention relates to an isolated breast tumour polynucleotide and the

CC polypeptide it encodes. The polynucleotide and polypeptide are useful for

CC detecting the presence of breast cancer in a patient, and in

CC pharmaceutical compositions for treating breast cancer. The sequences are

CC useful for stimulating an immune response in a patient and can therefore

CC be used in production of vaccines. The sequences are also useful for

CC detecting the presence of a cancer in a patient, by obtaining a

CC biological sample from the patient, contacting the biological sample with

CC a composition of the invention and detecting the amount of polynucleotide

CC that hybridizes to the sample. This sequence represents a human breast

CC tumour polypeptide of the invention

XX

SQ Sequence 650 AA;

Query Match 98.6%; Score 2023; DB 5; Length 650;

Best Local Similarity 98.7%; Pred. No. 7.1e-150;

Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MQKSVPNKALELKNQETLRADAILPSSESKQDYBESSWDSSELCTVSKQDVCLPKAHQ 60

Db 253 MQKSVPNKALELKNQETLRADAILPSSESKQDYBESSWDSSELCTVSKQDVCLPKAHQ 312

Qy 61 KEIDKINGLEGSPVKDGLLKANCCKMKSIPPTKALELMDMOTFKAEPPKPSAFEPAIEM 120

Db 313 KEIDKINGLEGSPVKDGLLKANCCKMKSIPPTKALELMDMOTFKAEPPKPSAFEPAIEM 372

Qy 121 QKSVPNKALELKNQETLRADAILPSSESKQDYBESSWDSSELCTVSKQDVCLPKATHQK 180

Db 373 QKSVPNKALELKNQETLRADAILPSSESKQDYBESSWDSSELCTVSKQDVCLPKAXHQK 432

Qy 181 EIDKINGLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPPKPSAFEPAIEMQ 240

Db 433 EIDKINGLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPPKPSAFEPAIEMQ 492

Qy 241 KSVPNKALELKNQETLRADQMPFSESKQKXVENSWSDESILRETIVSKQVCVPKATHQKE 300

Db 493 KSVPNKALELKNQETLRADQMPFSESKQKXVENSWSDESILRETIVSKQVCVPKATHQKE 552

Qy 301 MDKISGKLEDSTSLKILDTIHSERARELQKHCEQRTGMEQMKKFCVLKKLSEAK 360

Db 553 MDKISGKLEDSTSLKILDTVHSCERARELQKHCEQRTGMEQMKKFCVLKKLSEAK 612

Qy 361 EIKSQLENOKVKEQELCSVRFLTLMMKLIISYMKIAC 398

Db 613 EIKSQLENOKVKEQELCSVRFLTLMMKLIISYMKIAC 650

RESULT 6

ABJ37736

ID ABJ37736 standard; protein; 650 AA.

XX

AC ABJ37736;

XX

DT 15-MAY-2003 (first entry)

XX

XX Human tumour-related protein - SEQ ID No 469.

DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;

XX tumour; breast cancer; cancer; immune response stimulation.

KW

XX Homo sapiens.

OS

XX WO200283956-A1.

PN

XX 24-OCT-2002.

PD

XX

XX 15-APR-2002; 2002WO-US012378.

PF

XX 13-APR-2001; 2001US-00834759.

PR

XX 07-DEC-2001; 2001US-00007805.

PR

XX 13-FEB-2002; 2002US-00076622.

PR

XX (CORI-) CORIXA CORP.

PA

XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;

XX Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;

PI Vedvick TS, McNeill PD, Durham M;

PI

XX WPI; 2003-103376/09.

DR

XX New polypeptide and polynucleotide useful for stimulating and/or

PT expanding T cells specific for a tumor protein and treating breast

PT cancer.

XX

XX Example 1; Page 291-293; 375pp; English.

PS

XX The invention comprises a method of stimulating and/or expanding T cells

CC specific for a tumour protein. The invention further comprises human

CC nucleic acids and proteins that are associated with tumours (e.g. breast

CC cancer). The method and sequences of the invention are useful for

CC stimulating and/or expanding T cells specific for a tumour protein,

CC detecting the presence of cancer, stimulating an immune response in a

CC patient and treating breast cancer. The present amino acid sequence

CC represents a human tumour-related protein

XX

SQ Sequence 650 AA;

Query Match 98.6%; Score 2023; DB 6; Length 650;
Best Local Similarity 98.7%; Pred. No. 7.1e-150;
Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MQKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKAAHQ 60
Db 253 MQKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKAAHQ 312

Qy 61 KEIDKINGKLEGSVPKDGILLKANCVMKVS IPTKALELMDMTFKAEPPKPSAFEPALIMQ 120
Db 313 KEIDKINGKLEGSVPKDGILLKANCVMKVS IPTKALELMDMTFKAEPPKPSAFEPALIMQ 372

Qy 121 QKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKATHQ 180
Db 373 QKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKATHQ 432

Qy 181 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 240
Db 433 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 492

Qy 241 KSPVNPKALELKNQETLRADQMPFSESKQDYESSWDSSELCTVSKQDVCPKATHQ 300
Db 493 KSPVNPKALELKNQETLRADQMPFSESKQDYESSWDSSELCTVSKQDVCPKATHQ 552

Qy 301 MDKISGKLEDSTLSKILDTIHS CERARELOKHCEQCTGRMEQMKKFCVLLKKLSEAK 360
Db 553 MDKISGKLEDSTLSKILDTIHS CERARELOKHCEQCTGRMEQMKKFCVLLKKLSEAK 612

Qy 361 EIKSQLENQKVKWEQELCSVRFLTLMKMKIISYMKIAC 398
Db 613 EIKSQLENQKVKWEQELCSVRFLTLMKMKIISYMKIAC 650

RESULT 7
ADL93131
ID ADL93131 standard; protein; 650 AA.
AC ADL93131;
XX
DT 20-MAY-2004 (first entry)
DE Human breast cancer-associated polypeptide #5.
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX Homo sapiens.
XX US2003166022-A1.
XX PD 04-SEP-2003.
XX PF 15-APR-2002; 2002US-00124805.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.

DR N-PSDB; ADL93125.
XX An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX Example 1; SEQ ID NO 469; 294pp; English.
XX The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents the amino acid sequence of a human breast cancer-
CC associated polypeptide.
XX

SQ Sequence 650 AA;

Query Match 98.6%; Score 2023; DB 7; Length 650;
Best Local Similarity 98.7%; Pred. No. 7.1e-150;
Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MQKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKAAHQ 60
Db 253 MQKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKAAHQ 312

Qy 61 KEIDKINGKLEGSVPKDGILLKANCVMKVS IPTKALELMDMTFKAEPPKPSAFEPALIMQ 120
Db 313 KEIDKINGKLEGSVPKDGILLKANCVMKVS IPTKALELMDMTFKAEPPKPSAFEPALIMQ 372

Qy 121 QKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKATHQ 180
Db 373 QKSPVNPKALELKNQETLRADDEILPSESKQDYESSWDSSELCTVSKQDVCLPKATHQ 432

Qy 181 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 240
Db 433 EIDKINGKLESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 492

Qy 241 KSPVNPKALELKNQETLRADQMPFSESKQDYESSWDSSELCTVSKQDVCPKATHQ 300
Db 493 KSPVNPKALELKNQETLRADQMPFSESKQDYESSWDSSELCTVSKQDVCPKATHQ 552

Qy 301 MDKISGKLEDSTLSKILDTIHS CERARELOKHCEQCTGRMEQMKKFCVLLKKLSEAK 360
Db 553 MDKISGKLEDSTLSKILDTIHS CERARELOKHCEQCTGRMEQMKKFCVLLKKLSEAK 612

Qy 361 EIKSQLENQKVKWEQELCSVRFLTLMKMKIISYMKIAC 398
Db 613 EIKSQLENQKVKWEQELCSVRFLTLMKMKIISYMKIAC 650

RESULT 8
ADE44421
ID ADE44421 standard; protein; 650 AA.
XX
AC ADE44421;
XX
DT 29-JAN-2004 (first entry)
DE Human breast cancer protein #5.
XX
KW human; breast tumour; cancer; vaccine; T cell stimulator;
XX T cell expander.
XX Homo sapiens.
XX US2003104366-A1.
XX PN 05-JUN-2003.
XX PD 17-APR-2000; 2000US-00551621.
XX PF 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.

```
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
XX
XX (JIANG/) JIANG Y.
PA (DILL/) DILLON D C.
PA (MITC/) MITCHAM J L.
PA (XUJ/) XU J.
PA (HARL/) HARLOCKER S L.
XX
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX WPI; 2004-020270/02.
DR N-PSDB; ADE44415.
XX
XX Novel isolated polypeptide comprising immunogenic portion of breast tumor
PT protein or its variant, useful for formulating vaccines for inhibiting
PT cancer development in a patient.
XX
XX Example 1; SEQ ID NO 469; 217pp; English.
XX
XX The invention relates to an isolated polypeptide comprising at least an
CC immunogenic portion of a breast tumour protein. The polynucleotide, its
CC polypeptide, its antibody, a pharmaceutical composition comprising the
CC fusion protein or the polynucleotide encoding it, a vaccine comprising
CC the fusion protein or the polynucleotide encoding it, an isolated T cell
CC population comprising T cells specific for a breast tumour protein, and a
CC method for removing tumour cells from a biological sample is useful for
CC inhibiting the development of a cancer in a patient. The polypeptide is
CC useful for stimulating and/or expanding T cells specific for a breast
CC tumour protein. Stimulating and/or expanding T cells specific for a
CC breast tumour protein is useful for inhibiting the development of a
CC cancer in a patient. The method additionally involves the step of cloning
CC at least one proliferated cell and then administering the cloned T cells
CC to the patient. The present sequence represents a breast cancer protein.
XX
XX Sequence 650 AA;
XX
XX Query Match 98.6%; Score 2023; DB 8; Length 650;
XX Best Local Similarity 98.7%; Pred. No. 7.1e-150;
XX Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
Qy 1 MQKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKAAHQ 60
Db 253 MQKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKAXHQ 312
Qy 61 KEIDKINGKLEGGSPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPKPSAFEPALIE 120
Db 313 KEIDKINGKLEGGSPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPKPSAFEPALIE 372
Qy 121 QKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKATHOK 180
Db 373 QKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKAXHQ 432
Qy 181 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPKPSAFEPALIE 240
Db 433 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKAEPPKPSAFEPALIE 492
Qy 241 KSVPNKALELKNQOTLRADQFPSESQKQNVENSWSSESLRETVSQKDVCPKATHOKE 300
Db 493 KSVPNKALELKNQOTLRADQFPSESQKQNVENSWSSESLRETVSQKDVCPKATHOKE 552
Qy 301 MDKISGKLEDSTLSKILDTHTSCERARELQKHCEQCTGMEQMKKFCVLKXKXSEAK 360
Db 553 MDKISGKLEDSTLSKILDTHTSCERARELQKHCEQCTGMEQMKKFCVLKXKXSEAK 612
Qy 361 EIKSQLENQKVKWQELCSVRFLTLMKMKIISYMKIAC 398
Db 613 EIKSQLENQKVKWQELCSVRFLTLMKMKIISYMKIAC 650
XX
XX RESULT 9
XX ID AAU33358
XX AAU33358 standard; protein; 743 AA.
```

```
XX AAU33358;
XX
XX 18-DEC-2001 (first entry)
XX Human breast cancer protein B726P fusion protein #2.
XX Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX gene therapy; immunogen.
XX Homo sapiens.
XX WO200179286-A2.
XX
XX 25-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US012164.
XX
XX 17-APR-2000; 2000US-00551621.
XX 08-JUN-2000; 2000US-00590751.
XX 22-JUN-2000; 2000US-00604287.
XX 20-JUL-2000; 2000US-00620405.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
XX N-PSDB; AAS47422.
XX
XX Breast Tumor Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer.
XX
XX Claim 23; Page 295-296; 297pp; English.
XX
XX The invention relates to isolated breast tumour proteins and nucleic
XX acids that encode them, including immunogenic fragments of the proteins.
XX Also included are expression vectors expressing the proteins, transformed
XX cells and antibodies raised against the proteins or an antigen presenting
XX cell expressing the protein. The proteins and nucleic acids may be used
XX in the prevention, diagnosis and treatment of diseases associated with
XX inappropriate breast tumour protein expression, i.e. breast tumours and
XX breast cancer e.g by gene therapy. The nucleic acids and their
XX complements may also be used as DNA probes in diagnostic assays to detect
XX and quantitate the presence of similar nucleic acids in samples, and
XX therefore which patients may be in need of restorative therapy. The
XX proteins, nucleic acids and antibodies may be used in assays to identify
XX modulators (e.g. antagonists) of breast tumour protein expression and
XX activity. The antibodies and antagonists may also be used to down
XX regulate expression and activity. The antibodies may also be used as
XX diagnostic agents for detecting the presence of the proteins in samples
XX (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
XX purification diagnostic techniques. The present sequence is a breast
XX tumour protein encoded by a cDNA from a breast tumour cDNA library
XX isolated by subtractive hybridisation against a normal breast cDNA
XX library
XX
XX Sequence 743 AA;
XX
XX Query Match 98.6%; Score 2023; DB 4; Length 743;
XX Best Local Similarity 98.7%; Pred. No. 8.6e-150;
XX Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX
Qy 1 MQKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKAAHQ 60
Db 346 MQKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKAXHQ 405
Qy 61 KEIDKINGKLEGGSPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPKPSAFEPALIE 120
Db 406 KEIDKINGKLEGGSPVKDGLLKANGCMKVSIPTKALELMDMQTFKAEPPKPSAFEPALIE 465
Qy 121 QKSVPNKALELKNQOTLRADIELPSESQKDYESSWDSSESLCETVSQKDVCLPKATHOK 180
XX
XX AAU33358 standard; protein; 743 AA.
```

Db 466 QKSVNKALELNKNEQTLRADEILPSESQKQDYESSWDSSESLCETVSQKDVCLPKAXHQ 525
Qy 181 BIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 240
Db 526 BIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 585
Qy 241 KSVNKALELNKNEQTLRADQMPFSESQKQYVENSWSSESLRETYSQKDVCPKATHQKE 300
Db 586 KSVNKALELNKNEQTLRADQMPFSESQKQYVENSWSSESLRETYSQKDVCPKATHQKE 645
Qy 301 MDKISGKLEEDSTLSKILDTIHSCEARELOKHCEQCTGKMEQMKKFCVLKKLSEAK 360
Db 646 MDKISGKLEEDSTLSKILDTIHSCEARELOKHCEQCTGKMEQMKKFCVLKKLSEAK 705
Qy 361 EIKSQLENQKQVKEQELCSVRFLTLMKMKIISYMKIAC 398
Db 706 EIKSQLENQKQVKEQELCSVRFLTLMKMKIISYMKIAC 743

RESULT 10

ABG78925
ID ABG78925 standard; protein; 743 AA.

XX AC ABG78925;

XX DT 15-NOV-2002 (first entry)

XX DE Human breast tumour polypeptide #16.

XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.

XX OS Homo sapiens.

XX PN US2002085998-A1.

XX PD 04-JUL-2002.

XX PF 13-APR-2001; 2001US-00834759.

XX PR 28-DEC-1998; 98US-00222575.

XX PR 02-APR-1999; 99US-00285480.

XX PR 23-JUN-1999; 99US-00339338.

XX PR 02-SEP-1999; 99US-00389681.

XX PR 03-NOV-1999; 99US-00433826.

XX PR 17-APR-2000; 2000US-00551621.

XX PR 08-JUN-2000; 2000US-00590751.

XX PR 22-JUN-2000; 2000US-00604287.

XX PR 20-JUL-2000; 2000US-00620405.

XX PA (CORI-) CORIXA CORP.

XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;

XX PI Henderson RA;

XX DR WPI; 2002-635657/68.

XX DR N-PSDB; ABS64023.

XX XX

XX PT Novel breast cancer polynucleotides and polypeptides encoded by the

XX PT polynucleotides, useful for detecting the presence of breast cancer in a

XX PT patient, and in pharmaceutical compositions, for treating breast cancer.

XX PS Disclosure; Page 226-227; 247pp; English.

XX CC The invention relates to an isolated breast tumour polynucleotide and the

CC tumour polypeptide of the invention
XX SQ Sequence 743 AA;
Query Match 98.6%; Score 2023; DB 5; Length 743;
Best local similarity 98.7%; Pred. No. 8.6e-150;
Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 MQKSVNKALELNKNEQTLRADEILPSESQKQDYESSWDSSESLCETVSQKDVCLPKAXHQ 60
Db 346 MQKSVNKALELNKNEQTLRADEILPSESQKQDYESSWDSSESLCETVSQKDVCLPKAXHQ 405
Qy 61 KEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 120
Db 406 KEIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 465
Qy 121 QKSVNKALELNKNEQTLRADEILPSESQKQDYESSWDSSESLCETVSQKDVCLPKATHQK 180
Db 466 QKSVNKALELNKNEQTLRADEILPSESQKQDYESSWDSSESLCETVSQKDVCLPKAXHQ 525
Qy 181 BIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 240
Db 526 BIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMTFKAEPPKPSAFEPALIMQ 585
Qy 241 KSVNKALELNKNEQTLRADQMPFSESQKQYVENSWSSESLRETYSQKDVCPKATHQKE 300
Db 586 KSVNKALELNKNEQTLRADQMPFSESQKQYVENSWSSESLRETYSQKDVCPKATHQKE 645
Qy 301 MDKISGKLEEDSTLSKILDTIHSCEARELOKHCEQCTGKMEQMKKFCVLKKLSEAK 360
Db 646 MDKISGKLEEDSTLSKILDTIHSCEARELOKHCEQCTGKMEQMKKFCVLKKLSEAK 705
Qy 361 EIKSQLENQKQVKEQELCSVRFLTLMKMKIISYMKIAC 398
Db 706 EIKSQLENQKQVKEQELCSVRFLTLMKMKIISYMKIAC 743

RESULT 11

ABJ37748
ID ABJ37748 standard; protein; 743 AA.

XX AC ABJ37748;

XX DT 15-MAY-2003 (first entry)

XX DE Human tumour-related protein - SEQ ID No 494.

XX KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;

XX KW tumour; breast cancer; cancer; immune response stimulation.

XX OS Homo sapiens.

XX PN WO200283956-A1.

XX PD 24-OCT-2002.

XX PF 15-APR-2002; 2002WO-US012378.

XX PR 13-APR-2001; 2001US-00834759.

XX PR 07-DEC-2001; 2001US-00007805.

XX PR 13-FEB-2002; 2002US-00076622.

XX XX (CORI-) CORIXA CORP.

XX PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;

XX PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;

XX PI Vedwick TS, McNeill PD, Durham M;

XX DR WPI; 2003-1033376/09.

XX XX New polypeptide and polynucleotide useful for stimulating and/or

XX PT expanding T cells specific for a tumor protein and treating breast

XX PT cancer.

XX Disclosure; Page 312-314; 375pp; English.

PS The invention comprises a method of stimulating and/or expanding T cells

XX specific for a tumour protein. The invention further comprises human

CC nucleic acids and proteins that are associated with tumours (e.g. breast

CC cancer). The method and sequences of the invention are useful for

CC stimulating and/or expanding T cells specific for a tumour protein,

CC detecting the presence of cancer, stimulating an immune response in a

CC patient and treating breast cancer. The present amino acid sequence

CC represents a human tumour-related protein

XX Sequence 743 AA;

Query Match 98.6%; Score 2023; DB 6; Length 743;

Best Local Similarity 98.7%; Pred. No. 8.6e-150;

Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKAAHQ 60

DB 346 MQKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKAAHQ 405

QY 61 KEIDKINGKLEGSVPKDGLLKANGCMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 120

DB 406 KEIDKINGKLEGSVPKDGLLKANGCMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 465

QY 121 QKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKATHQ 180

DB 466 QKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKATHQ 525

QY 181 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 240

DB 526 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 585

QY 241 KSVPNKALELKNQETLRADQMPFSESQKQKXVEENSWDSSESLRETVSQKDVCPKATHQ 300

DB 586 KSVPNKALELKNQETLRADQMPFSESQKQKXVEENSWDSSESLRETVSQKDVCPKATHQ 645

QY 301 MDKISGLESTSLSKILDTIHSCERARELOKHCEQCTGMEQMKKFCVLKKKLEAK 360

DB 646 MDKISGLESTSLSKILDTIHSCERARELOKHCEQCTGMEQMKKFCVLKKKLEAK 705

QY 361 EIKSQLENQKQKWEQELCSVRFLTLMKMKIISYMKIAC 398

DB 706 EIKSQLENQKQKWEQELCSVRFLTLMKMKIISYMKIAC 743

RESULT 12

ADL93156

ID ADL93156 standard; protein; 743 AA.

XX ADL93156;

XX 20-MAY-2004 (first entry)

DT Human breast cancer-associated polypeptide fusion protein #2.

DE gene therapy; protein therapy; vaccine; breast cancer; cancer; human.

XX Homo sapiens.

OS US2003166022-A1.

PN 04-SEP-2003.

PD 15-APR-2002; 2002US-00124805.

PP 28-DEC-1998; 98US-00222575.

XX 02-APR-1999; 99US-00285480.

PR 23-JUN-1999; 99US-00339338.

PR 02-SEP-1999; 99US-00389681.

PR 03-NOV-1999; 99US-00433826.

PR 17-APR-2000; 2000US-00551621.

PR 08-JUN-2000; 2000US-00590751.

PR 22-JUN-2000; 2000US-00604287.

PR 20-JUL-2000; 2000US-00620405.

PR 13-APR-2001; 2001US-00834759.

PR 07-DEC-2001; 2001US-00007805.

PR 13-FEB-2002; 2002US-00076622.

XX (CORI-) CORIXA CORP.

XX Houghton RL, Sleath PR, Persing DH;

XX WPI; 2003-874918/81.

DR N-PSDB; ADL93153.

XX An isolated oncogenic polypeptide useful for preventing, diagnosing and

PT treating breast cancer.

PT Disclosure; SEQ ID NO 494; 294pp; English.

PS The invention relates to an isolated breast cancer-associated

CC polypeptide. The polypeptide may be used for the diagnosis and treatment

CC of breast cancers. The methods are useful for detecting the presence of a

CC cancer in a patient and treating a cancer in a patient. The present

CC sequence represents the amino acid sequence of a human breast cancer-

CC associated polypeptide.

XX Sequence 743 AA;

Query Match 98.6%; Score 2023; DB 7; Length 743;

Best Local Similarity 98.7%; Pred. No. 8.6e-150;

Matches 393; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MQKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKAAHQ 60

DB 346 MQKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKAAHQ 405

QY 61 KEIDKINGKLEGSVPKDGLLKANGCMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 120

DB 406 KEIDKINGKLEGSVPKDGLLKANGCMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 465

QY 121 QKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKATHQ 180

DB 466 QKSVPNKALELKNQETLRADDEILPSESQKDYESSWDSSESLCETVSQKDVCLPKATHQ 525

QY 181 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 240

DB 526 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMOTFKAEPEKPSAFEPAIEM 585

QY 241 KSVPNKALELKNQETLRADQMPFSESQKQKXVEENSWDSSESLRETVSQKDVCPKATHQ 300

DB 586 KSVPNKALELKNQETLRADQMPFSESQKQKXVEENSWDSSESLRETVSQKDVCPKATHQ 645

QY 301 MDKISGLESTSLSKILDTIHSCERARELOKHCEQCTGMEQMKKFCVLKKKLEAK 360

DB 646 MDKISGLESTSLSKILDTIHSCERARELOKHCEQCTGMEQMKKFCVLKKKLEAK 705

QY 361 EIKSQLENQKQKWEQELCSVRFLTLMKMKIISYMKIAC 398

DB 706 EIKSQLENQKQKWEQELCSVRFLTLMKMKIISYMKIAC 743

RESULT 13

ABJ37782

ID ABJ37782 standard; protein; 661 AA.

XX ABJ37782;

XX 15-MAY-2003 (first entry)

DT Human tumour-related protein - SEQ ID No 552.

DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;

XX tumour; breast cancer; cancer; immune response stimulation.

OS Homo sapiens.
XX WO2000283956-A1.
XX 24-OCT-2002.
XX 15-APR-2002; 2002WO-US012378.
XX 13-APR-2001; 2001US-00834759.
XX 07-DEC-2001; 2001US-00007805.
XX 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX Example 8; Page 332-334; 375pp; English.
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumor protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumors (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumor protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX
SQ Sequence 661 AA;
Query Match 98.3%; Score 2018; DB 6; Length 661;
Best Local Similarity 98.2%; Pred. No. 1.8e-149; Indels 0; Gaps 0;
Matches 391; Conservative 4; Mismatches 3;
QY 1 MOKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKAAHQ 60
DB 264 MOKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKAAHQ 323
QY 61 KEIDKINGKLEGGSPVKDGLLKANCCKMVSIPTKALELMDMTFKABPEPKPSAFPAIEM 120
DB 324 KEIDKINGKLEGGSPVKDGLLKANCCKMVSIPTKALELMDMTFKABPEPKPSAFPAIEM 383
QY 121 OKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKATHQK 180
DB 384 OKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKAAHQ 443
QY 181 EIDKINGKLEGGSPNDGFLKAPCRMKVSIPTKALELMDMTFKABPPKPSAFPAIEMQ 240
DB 444 EIDKINGKLEGGSPNDGFLKAPCRMKVSIPTKALELMDMTFKABPPKPSAFPAIEMQ 503
QY 241 KSVNKALELNKNEQTLRADQMPFSSKQKNVENSNDSESLRETYSQKDVCPKATHQKE 300
DB 504 KSVNKALELNKNEQTLRADQMPFSSKQKNVENSNDSESLRETYSQKDVCPKATHQKE 563
QY 301 MDKISGKLEDSTLSKILDTIHSCEARELOKHCEQCTGKMEQKKFCVLKXKLEAK 360
DB 564 MDKISGKLEDSTLSKILDTIHSCEARELOKHCEQCTGKMEQKKFCVLKXKLEAK 623
QY 361 EIKSOLENKQVKEOBLCSVRFLTLMMKIIISYMKIAC 398
DB 624 EIKSOLENKQVKEOBLCSVRFLTLMMKIIISYMKIAC 661

ID ADL93214 standard; protein; 661 AA.
XX
AC ADL93214;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human breast cancer-associated polypeptide #47.
XX gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX Homo sapiens.
OS
XX US2003166022-A1.
XX
XX 04-SEP-2003.
XX
XX 15-APR-2002; 2002US-00124805.
XX
XX 28-DEC-1998; 98US-00222575.
XX 02-APR-1999; 99US-00285480.
XX 23-JUN-1999; 99US-00339338.
XX 02-SEP-1999; 99US-00389681.
XX 03-NOV-1999; 99US-00433826.
XX 17-APR-2000; 2000US-00551821.
XX 08-JUN-2000; 2000US-00590751.
XX 22-JUN-2000; 2000US-00604287.
XX 20-JUL-2000; 2000US-00620405.
XX 13-APR-2001; 2001US-00834759.
XX 07-DEC-2001; 2001US-00007805.
XX 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX
XX Houghton RL, Sleath PR, Persing DH;
PI
XX WPI; 2003-874918/81.
XX N-PSDB; ADL93211.
XX
XX An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX
XX Example 8; SEQ ID NO 552; 294pp; English.
XX
XX The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents the amino acid sequence of a human breast cancer-
CC associated polypeptide.
XX
SQ Sequence 661 AA;
Query Match 98.3%; Score 2018; DB 7; Length 661;
Best Local Similarity 98.2%; Pred. No. 1.8e-149; Indels 0; Gaps 0;
Matches 391; Conservative 4; Mismatches 3;
QY 1 MOKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKAAHQ 60
DB 264 MOKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKAAHQ 323
QY 61 KEIDKINGKLEGGSPVKDGLLKANCCKMVSIPTKALELMDMTFKABPEPKPSAFPAIEM 120
DB 324 KEIDKINGKLEGGSPVKDGLLKANCCKMVSIPTKALELMDMTFKABPEPKPSAFPAIEM 383
QY 121 OKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKATHQK 180
DB 384 OKSVNKALELNKNEQTLRADEILPSESKQDYESSWDSLSCLCTVTSQKDVCLPKAAHQ 443
QY 181 EIDKINGKLEGGSPNDGFLKAPCRMKVSIPTKALELMDMTFKABPPKPSAFPAIEMQ 240
DB 444 EIDKINGKLEGGSPNDGFLKAPCRMKVSIPTKALELMDMTFKABPPKPSAFPAIEMQ 503
QY 241 KSVNKALELNKNEQTLRADQMPFSSKQKNVENSNDSESLRETYSQKDVCPKATHQKE 300

Db 504 KSPVKNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKATHQKE 563
Qy 301 MDKISGKLEDSTSLSKILDTIHSCERARELQKHCEQCTGMEQMKKKFCVLKKGKJSEAK 360
Db 564 MDKISGKLEDSTSLSKILDTVHSCERARELQKHCEQCTGMEQMKKKFCVLKKGKJSEAK 623
Qy 361 EIKSQLENQKVKEQELCSVRFLTLMKWKIISYMKIAC 398
Db 624 EIKSQLENQKVKEQELCSVRFLTLMKWKIISYMKIAC 661

RESULT 15

AAU33351
ID AAU33351 standard; protein; 1002 AA.
XX AC AAU33351;
XX AC
DT 18-DEC-2001 (first entry)
XX DE Human breast cancer protein B726P from alternatively spliced cDNA.
XX KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
XX KW gene therapy; immunogen.
XX OS Homo sapiens.
XX XX
XX FN WO200179286-A2.
XX XX
XX PD 25-OCT-2001.
XX XX
XX PF 12-APR-2001; 2001WO-US012164.
XX XX
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX XX
XX DR WPI; 2001-611721/70.
XX DR N-PSDB; AAS47411.
XX XX
XX PT Breast Tumor Proteins and nucleic acids useful for the prevention,
XX PT diagnosis and treatment of breast cancer.
XX XX
XX PS Claim 3; Page 281-283; 297pp; English.
XX XX

CC The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
XX library

SQ Sequence 1002 AA;

Query Match 94.6%; Score 1940.5; DB 4; Length 1002;
Best Local Similarity 98.4%; Pred. No. 3.8e-143;
Matches 379; Conservative 1; Mismatches 4; Indels 1; Gaps 1;
Qy 1 MOKSVPNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKAAHQ 60
Db 253 MOKSVPNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKAAHQ 312
Qy 61 KEIDKINGKLEGSFVKDGLLKANCCKMKSIPPTKALELMDMQTFKABPPEKPSAFEPAIEM 120
Db 313 KEIDKINGKLEGSFVKDGLLKANCCKMKSIPPTKALELMDMQTFKABPPEKPSAFEPAIEM 372
Qy 121 OKSVPNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKATHQK 180
Db 373 OKSVPNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKATHQK 432
Qy 181 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKABPPEKPSAFEPAIEMQ 240
Db 433 EIDKINGKLEESPDNDGFLKAPCRMKVSIPTKALELMDMQTFKABPPEKPSAFEPAIEMQ 492
Qy 241 KSVPNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKATHQKE 300
Db 493 KSVPNKALELKNQETLRADQMFSESKQKVENSWSSESLRETYSQKDVCPKATHQKE 552
Qy 301 MDKISGKLEDSTSLSKILDTIHSCERARELQKHCEQCTGMEQMKKKFCVLKKGKJSEAK 360
Db 553 MDKISGKLEDSTSLSKILDTVHSCERARELQKHCEQCTGMEQMKKKFCVLKKGKJSEAK 612
Qy 361 EIKSQLENQKVKEQELCSVRFLTL 385
Db 613 EIKSQLENQKVKEQELCSVR-LTL 636

Search completed: August 1, 2005, 22:18:03
Job time : 144.469 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:08:06 ; Search time 24.5213 Seconds
(without alignments)
1561.672 Million cell updates/sec

Title: US-09-489-079-24
Perfect score: 2052
Sequence: 1 MQKSPNKALEKNEQTLRA.....SVREFTLMKMKIISYMKIAC 398

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	182.5	8.9	2954	T14156	kinesin-related pr
2	182	8.9	3488	T34418	hypothetical prote
3	179.5	8.7	1313	F96673	hypothetical prote
4	173	8.4	585	A24168	involucrin - huma
5	171.5	8.4	2663	E89066	centromere protei
6	169	8.2	2109	T33247	protein H05O09.1
7	169	8.2	2109	T33247	hypothetical prote
8	168	8.2	1356	S32763	kinectin 1 - huma
9	165.5	8.1	853	T51505	hypothetical prote
10	164.5	8.0	993	A49461	synaptonemal compl
11	164.5	8.0	1690	T13030	microtubule bindin
12	163	7.9	1269	F84730	probable myosin he
13	163	7.9	2116	A26655	myosin heavy chain
14	162.5	7.9	1938	T49464	alpha cardiac myos
15	162.5	7.9	2245	T18278	myosin heavy chain
16	162	7.9	1939	A46762	myosin alpha heavy
17	161.5	7.9	944	S26710	spindle pole body
18	161.5	7.9	1940	A24922	myosin heavy chain
19	161.5	7.9	2020	T21174	hypothetical prote
20	161	7.8	1937	T38055	myosin heavy chain
21	158.5	7.7	1738	T14867	interaptin - slime
22	158.5	7.7	1939	T148175	myosin heavy chain
23	158.5	7.7	1940	S04090	myosin heavy chain
24	158	7.7	1938	SC5421	smooth muscle myos
25	158	7.7	1972	JC5420	smooth muscle myos
26	156.5	7.6	1392	A43336	microtubule-vesicl
27	156.5	7.6	1427	S22695	restin - human
28	156.5	7.6	1938	S06005	myosin alpha heavy
29	156	7.6	1133	T22976	hypothetical prote

30	155.5	7.6	474	2	S39475	embryonic protein
31	155.5	7.6	1534	2	A56734	ribosome receptor,
32	155.5	7.6	1804	2	T34518	nestin - golden ha
33	155	7.6	3187	2	JC5837	364k Golgi complex
34	154.5	7.5	635	1	I37060	involucrin L - gor
35	154.5	7.5	750	2	T38435	coiled coil protei
36	154	7.5	1790	2	S67593	transport protein
37	153.5	7.5	1164	2	T24806	hypothetical prote
38	153.5	7.5	1964	2	A59282	nonmuscle myosin I
39	153.5	7.5	6642	2	T29757	protein UNC-89 - C
40	152.5	7.4	841	2	A86188	hypothetical prote
41	152.5	7.4	1300	2	I53799	CGI protein - huma
42	152.5	7.4	2139	2	T18296	myosin heavy chain
43	152.5	7.4	2331	2	T25410	hypothetical prote
44	152	7.4	624	2	PC6003	surface membrane p
45	152	7.4	978	2	A70387	conserved hypothet

ALIGNMENTS

RESULT 1

T14156
kinesin-related protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14156
R:Wood, K.W.; Sakowicz, R.; Goldstein, L.S.; Cleveland, D.W.
Cell 91, 357-366, 1997
A>Title: CNP-E is a plus end-directed kinetochore motor required for metaphase chromosome
A:Reference number: Z17893; MUID:98028574; PMID:9363944
A:Accession: T14156
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2954 <WOO>
A:Cross-references: UNIPROT:O42263; EMBL:AF027728; NID:g2586070; PID:g2586071; PIDN:AAC66
C:Genetics:
A:Gene: XCENP-E
C:Superfamily: centromere protein E; kinesin motor domain homology

Query Match 8.9%; Score 182.5; DB 2; Length 2954;

Best Local Similarity 22.6%; Pred. No. 0.026; Mismatches 171; Indels 113; Gaps 21;

Matches 107; Conservative 83;	11	ELKNEQ-TLRADILPSESQKDYESSWDSLSCE	TVSQKDVCLPKAAHQEIDKIN--	67
	1792	ELKNSQRTVIAER----	DQLQDDLRES---VEMSIET--QDDLKRAQEA	LQKKQKQVQL 1842
QY	68	-----GKLEGSV-----	KDGLLKANCCKMKVSIPTKALELMDMOTPK	104
DB	1843	TSQISVLQEKISLLENQMLYNVATVKTLSER	DDLNSQKHLPSEITLSLSLKEKB-FA	1901
QY	105	AEPPF--KPSAFEPALPMQKSPNKALEKNEQ	TLRADILPSESKQDYESSWDSLSL	162
DB	1902	LEQAERKADAAKTIDITEKISNIEEQLQATN	LKETLYERESLIQCKEQLALNTEHL	1961
QY	163	CETVSQKDVCLPKAQHKE-----	IDKINGKLESPDNDG-----FLKA	201
DB	1962	RETLKSKDLALGKMEQERDEANKVIALTEK	MSLSEEQINENVTTLKEGEKETFYLR	2021
QY	202	PCR-----MKVSIPTKALELMDMOTFKAP	PE-----KPSAFEPALPMQKSPN	245
DB	2022	PSKQSSSQMEELRESLTKDLQLEAEKEI	SEATNEIKNLTKAKISLSEEEILQNASILN	2081
QY	246	KALELKNEQTLRADQMPSESQKQVENSWS	SESLRETVSQKDVCLPKAQHKE--MDK	303
DB	2082	EA--VSERENLR-----HSKQQLVSE---	LEQLSLTLKSRDHAFQSKREKEDAVNK	2128
QY	304	ISGKLEDSTSLKILDTIHSCEARELEQKH	CH--EQCTGKME-----QMK-----	346
DB	2129	IASLAEIKILTKENWDEPDFDSKESLQES	SHLSEELCTYKTELQMLKQCKEDINNKLA	2188
QY	347	-KKFCVLKKLSEAKE----	IKSQLENQKVKEQELCSVRFLTLMMOKIIS	YNK 395

A:Reference number: A57786; MUID:89053976; PMID:2461365
A:Contents: annotation; transglutaminase-catalyzed cross-link sites of intact and fragmented
A>Note: in vitro studies of native, soluble involucrin showed almost exclusive preferential
er modifications in vivo, allows modification at a number of other sites
C:Comment: During the terminal differentiation of keratinocytes, this protein from the c
linked envelope under the plasma membrane.

C:Genetics:
A:Gene: GDB:IVL
A:Cross-references: GDB:119355; OMIM:147360
A:Map position: 1q21-1q21
C:Superfamily: involucrin
C:Keywords: cornified cell envelope; duplication; epidermis; tandem repeat
F:153-541/Region: 10-residue repeats (Q-E-G-Q-[PIV]-[KE]-[LH]-[PL]-E-Q)
F:496/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of other proteins) #statu

Query Match 8.4%; Score 173; DB 1; Length 585;
Best Local Similarity 23.1%; Pred. No. 0.014; Mismatches 160; Indels 70; Gaps 16;
Matches 92; Conservative 70;
Db 7 NKALBLKNEOTLRADRI---LPSESKQKDY-----ESSWDSLSLCTVSKQDVC 53
103 NPEQQLKQEKTDQDQQLNKLSEKLLDQQLDQLVKEDEQLGMKKEQLLELPEQOE-- 160
QY 54 LPKAAHQEKIDKINGKLESPVQDGLLKANGCMKVISIPTK---ALELMDMTFKAPPEPK 110
Db 161 ----GHLKHLQEQGQKHPQEQGQLE-----LPQEQGQLELPEQEQGQLELPEQ 208
QY 111 PSAPFATIMQKSVNPK---ALELNEOTLRADRIILPSESKQDYEESSWDSLSLCTVTS 167
Db 209 Q-----EGQLELPEQEQGQLELPEQEQ-----EGQLELSEQEQGQLELSEQEQGQLELSE 257
QY 168 QKDVLCPKATHQEKIDKINGKLESPDNDGFLKPCRMKVISIPTKALELMDMTFKABPP 227
Db 258 QEQGQLKHLQEQ-----GQLEVPPEQMGQLK-----YLEQEQGQLKHLQEQEQKPELP 306
QY 228 EKPSAPFAPAIEMQKSVNPKALBLKNEOTLRADMTPEPSESKQKQNVENSWDSLSLRETVSQ 287
Db 307 EQMGQQLKHLQEQEQGP-KHLE---QEQGQLELSEQEQGQLELSEQEQGQLELSEHLEH 362
QY 288 KDVCVPKATHQEMDKISK---LEDSTSLSKILDTHSCERARELQKDCFCQCTGKMBQ 344
Db 363 LGLPEQQLVQLKLEKQKQKHLSEEGQLK-----HLVQEQGQLK---HLVQEQGQLEQ 415
QY 345 MKKFCVLKKLSEAKEIKSLENQKVKWEQE 376
Db 416 QERQVEHLEQVQQLKHLSEQ-EGQLKHLQEQ 446

RESULT 5

S28261
centromere protein E - human
N:Alternate names: centromere 312K protein; kinesin-related protein CENP-E
C:Species: Homo sapiens (man)
C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 09-Jul-2004
C:Accession: S28261
R:Yen, T.J.; Li, G.; Schaar, B.T.; Szilak, I.; Cleveland, D.W.
Nature 359, 536-539, 1992
A:Title: CENP-E is a putative kinetochore motor that accumulates just before mitosis.
A:Reference number: S28261; MUID:93024922; PMID:1406971
A:Accession: S28261
A:Molecule type: mRNA
A:Residues: 1-2663 <YEN>
A:Cross-references: UNIPROT:Q02224; EMBL:Z15005; NID:g29864; PIDN:CAA78727.1; PID:g29865
C:Genetics:
A:Gene: GDB:CENPE
A:Cross-references: GDB:361164; OMIM:117143
A:Map position: 4q24-4q25
C:Superfamily: centromere protein E; kinesin motor domain homology
C:Keywords: ATP; coiled coil; microtubule binding; mitosis; nucleotide binding; P-loop
F:7-335/Domain: kinesin motor domain homology <KMT>
F:86-93/Region: nucleotide-binding motif A (P-loop)
F:486-2183/Domain: coiled coil #status predicted <COI>
F:92/Binding site: ATP (Lys) #status predicted

Query Match 8.4%; Score 171.5; DB 1; Length 2663;
Best Local Similarity 21.4%; Pred. No. 0.095;
Matches 103; Conservative 84; Mismatches 157; Indels 137; Gaps 24;
QY 1 MOKSVPNKALELKNQOTLRADRIILPSESKQKDY-----ESSWDSLSLCTVSKQDVC 54
Db 725 LQEL-NK--EVENEARREEVILLSELKSPSEVERLRKEIQDKSEELHITTSKDKLF 781
QY 55 PKAAHQEKIDKINGKLEG-SPVQDGLLKANGCMKVISIPTKALELMDMTFKAPPEPKPSA 113
Db 782 SEVVHKE--SRVQGLLEIGTKDQDLATQSNYKSTD-----QEQNFKTLHMDFEQK 832
QY 114 FEPAIE---MOKSVPNKALE-----LKNQOTLRADRIILPSESKQKDYEESSWDS 159
Db 833 YKVLLEENRMNQEIIVNLSEKAQKFDSSLGALKTELSYKTQEL---QEKTRVQERLNE 889
QY 160 ESLCTVSKQDVCPLKATHQEKI--DKINGKLE---SPDNDGF-----LK 200
Db 890 BOLKEQLENRDSPLQTVREKTLITEKLQOTLSEVKTITQEKDDLKQLQESLQIERDQLK 949
QY 201 APCR--MKVSIPTK-----ALELM-----DMQTFKAPPEPKPS-----AFEP 240
Db 950 SDIHDVTVMNIDTQQLRNALLESKQHOETINTLAKSISEEVRNLHMEENTGETKDEFQ 1009
QY 241 KSV-----PNKALELKNQOTLRAD-----DMQTFKAPPEPKPS-----QMFPS 268
Db 1010 QKMGIDKKQDLEAKNTQTLTADVKDNEIIEQQRKIFSLIOEKNELQQLMESVIAKEQL 1069
QY 269 -----KNVEENSWSESIR---ETVSQKDVCPKATHQEKMDKISGKLEDSTSLSKILD 319
Db 1070 KTDLENTEMTIENQELRLGLDELKQOEIVQAEKNH-----AKKEGE 1114
QY 320 THSCERARELQ---KDHCEQCTGKWEQ---KKKFCVLKKLSEAKEIKSLENQKVKW 373
Db 1115 LSRTCDRLAEVVEKLEKSKQQLQEQQLLVQVEEMSEMQKINIEINLKNELKNKELTL 1174
QY 374 E 374
Db 1175 E 1175

RESULT 6

E89066
protein H05009.1 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89066
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E89066
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2109 <STO>
A:Cross-references: UNIPROT:Q8ISF7; UNIPROT:Q8ISF6; UNIPROT:Q8ISF5; GB:chr_V; PIDN:AAAC19;
C:Genetics:
A:Gene: H05009.1
A:Map position: 5
Query Match 8.2%; Score 169; DB 2; Length 2109;
Best Local Similarity 23.1%; Pred. No. 0.1;
Matches 96; Conservative 77; Mismatches 173; Indels 70; Gaps 17;
QY 1 MOKSVPNKALELKNQOTLRADRIILPSESKQKDYEESSWDSLSLCTVSKQD---VCLPKA 57
Db 1592 IQQEVPTVAAPSEPTQADVPEKAPSEPSQADVPKVAAPLE---QTQIQQEVPMVAAPLE 1648
QY 58 AHQEKIDKINGKLESPVQDGLLKANGCMKVISIPTKALELMDMTFKAPPEPKPSAFEPA 117

Db 1649 PTQADVPKVAAPLEQSQIQ-----EVPTVA---APSEPTQADVPKAAPSEPS 1694
Qy 118 -----TEMQSVPNKALELKNQETLRADLILPSESQKQYEESSWDSLSICE 164
Db 1695 QADVPKVAAPLEQIQIQEVPMVAAPLEPIQIEVPKKAAPSEPTQEDVPKGAAPLEPTQE 1754
Qy 165 TVSQKDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVGIPTKALELMDMQTFKA 224
Db 1755 DVPE--AAPSPTQEDVPK-----EAPSEPTQEDVPKKAAPSEPTQ--ENVKKAAPS 1805
Qy 225 EP-PEKPSAFPAJEMQKSVPNKA-----LELKNQETLRADQMPFSESQKQNVENSWS 278
Db 1806 EPTKDVPEKAAPSEPIQIEVPKKAATLSPTQEQSEVSKRSPVPTQIQQAASEET-PL 1864
Qy 279 ESLRETYSQ-----KDVCPK-ATHQKEMDKISKLEDSTSL-----SKILDTIHSCE 325
Db 1865 EETNETVQTTNEDVKEAEPVNAEAQKVD--SSDLQVAASEIAHLAIDEAVLETSNQPS 1922
Qy 326 RARELQKDHCEQCTGKMEQMK--KKFCVLKKLSEAKBIKSQLENQKVMQBELCS 379
Db 1923 QFDSLQ-----EQKPSVHNEHVRSCVDLTFSRDSQIVSDVIAEVGDEDECS 1974

RESULT 7
T33247
hypothetical protein H05009.1 - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33247
R:Leonard, S.; Hinds, K.
A:Description: The sequence of C. elegans cosmid H05009.
A:Reference number: 221308
A:Accession: T33247
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2109 <LEO>
A:Cross-references: UNIPROT:Q81SF7; UNIPROT:Q81SF6; UNIPROT:Q81SF5; EMBL:AF067951; PIDN:
A:Experimental source: strain Bristol N2; clone H05009
C:Genetics:
A:Gene: CESP:H05009.1
A:Map position: 5
A:Note: intron positions not resolved (incomplete sequence)

Query Match 8.2%; Score 169; DB 2; Length 2109;
Best Local Similarity 23.1%; Pred. No. 0.1;
Matches 96; Conservative 77; Mismatches 173; Indels 70; Gaps 17;

Qy 1 MQKSVPNKALELKNQETLRADLILPSESQKQYEESSWDSLSICEVTSQKD---VCLPKA 57
Db 1592 IQQEVPTVAAPSEPTQADVPKKAAPSPSQADVPKVAAPLE---OTQIQEVPMVAAPLE 1648
Qy 58 AHQKEIDKINGKLESPVKDGLLKANCGMKVSIPTKALELMDMQTFKAPEKPSAFAPA 117
Db 1649 PTQADVPKVAAPLEQSQIQ-----EVPTVA---APSEPTQADVPKKAAPSEPS 1694
Qy 118 -----TEMQSVPNKALELKNQETLRADLILPSESQKQYEESSWDSLSICE 164
Db 1695 QADVPKVAAPLEQIQIQEVPMVAAPLEPIQIEVPKKAAPSEPTQEDVPKGAAPLEPTQE 1754
Qy 165 TVSQKDVCLPKATHQKEIDKINGKLESPDNDGFLKAPCRMKVGIPTKALELMDMQTFKA 224
Db 1755 DVPE--AAPSPTQEDVPK-----EAPSEPTQEDVPKKAAPSEPTQ--ENVKKAAPS 1805
Qy 225 EP-PEKPSAFPAJEMQKSVPNKA-----LELKNQETLRADQMPFSESQKQNVENSWS 278
Db 1806 EPTKDVPEKAAPSEPIQIEVPKKAATLSPTQEQSEVSKRSPVPTQIQQAASEET-PL 1864
Qy 279 ESLRETYSQ-----KDVCPK-ATHQKEMDKISKLEDSTSL-----SKILDTIHSCE 325
Db 1865 EETNETVQTTNEDVKEAEPVNAEAQKVD--SSDLQVAASEIAHLAIDEAVLETSNQPS 1922
Qy 326 RARELQKDHCEQCTGKMEQMK--KKFCVLKKLSEAKBIKSQLENQKVMQBELCS 379

Db 1923 QFDSLQ-----EQKPSVHNEHVRSCVDLTFSRDSQIVSDVIAEVGDEDECS 1974

RESULT 8
S32763
kinectin 1 - human
C:Species: Homo sapiens (man)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S32763; I37947
R:Kruppa, G.; Fuetterer, A.; Lemke, H.; Kroenke, M.
submitted to the EMBL Data Library, April 1993
A:Description: Cloning and characterization of TAF, a novel transactivating protein.
A:Reference number: S32763
A:Accession: S32763
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1356 <KRU>
A:Cross-references: UNIPROT:Q86UP2; EMBL:Z22551
R:Futterer, A.; Kruppa, G.; Kramer, B.; Lemke, H.; Kronke, M.
Mol. Biol. Cell 6, 161-170, 1995
A:Title: Molecular cloning and characterization of human kinectin.
A:Reference number: I37947; MUID:95306853; PMID:7787243
A:Accession: I37947
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1356 <RES>
A:Cross-references: EMBL:Z22551; NID:g296163; PIDN:CAA80271.1; PID:g296164
C:Genetics:
A:Gene: GDB:KTN1
A:Cross-references: GDB:6165852; OMIM:600653

Query Match 8.2%; Score 168; DB 2; Length 1356;
Best Local Similarity 22.5%; Pred. No. 0.068;
Matches 100; Conservative 83; Mismatches 153; Indels 108; Gaps 22;

Qy 1 MQKSVPNKALELKNQETLRADLILPSESQKQYEESSWDSLSICEVTSQKD---SOKDVC-- 53
Db 732 MEKCIQEKDEKLTVBELLTGLIQVATKEELNATPENSLLTKEVQDLKAKQNDQVSF 791
Qy 54 -----LPKAAHQKEIDKINGKLESPVKDGLLKANCGMKVSIPTKAL----- 95
Db 792 ASLVEELKVKIHEKD-----GKIKS---VBELLEAEL-LKVANKEKTVDLQKEIKALKE 842
Qy 96 ELMQDQTFKAPEKPSAFAPAEMQKSVPNKALELKN-----EOTLRADLILPSESQK 149
Db 843 EIGNVQLEKQA-----QLSITSKVQELQNLKKEEQMNTMKAVL--BEKE 886
Qy 150 KDYEES--SW-----DSSEL---CETVSKQDVCLPKATHQKEIDKINGKLESPDNDGFL 199
Db 887 KDLANTGKWLQDQEEENESLKAHVQEAQHN--LKEASSASQPEELEIVLKEKGNELKRL 944
Qy 200 KAPCRMKVS-IPTKALELMDQ-----TFFKAPEKPSAFAPAEMQKSVPNKA 247
Db 945 EAMLKERESDLSSKTQLLDVQVDENKLFQIEQLKQNYQQAASSFPFPHELLKVISERE 1004
Qy 248 LELK--NEOTLRADQMPFSESQKQNVENSWS--ESLRETYSQKDVCPKATHQKEMDK 303
Db 1005 KEISGLMNELDLSKDAVEHQRKNNDLRKNWAMEALATEKMLQDKVKNKTSKERQQOV 1064
Qy 304 ISGKLEDSTSLSKIL-----DTIHSCE-RARELQKDHCEQCTGKMEQMKKFC 350
Db 1065 EAVELEAKEVLLKLFPPKVSVPNSLYGEWLHGFEKKAKE-----CMAGTSGSEEVK---- 1115
Qy 351 VLKKKLSAEKBIKS--OLENQKVK 372
Db 1116 VLEHLKLEADEMHTLLOLECEKYK 1139

RESULT 9
T51505
hypothetical protein F5B19.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: T51505
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z55394
A:Accession: T51505
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-853 <SAT>
A:Cross-references: UNIPROT:Q9LPE4; EMBL:AL391147
A:Experimental source: cultivar Columbia; BAC clone F5E19
C:Genetics:
A:Map position: 5
A:Introns: 6/2; 79/3
A:Note: F5E19_70

Query Match 8.1%; Score 165.5; DB 2; Length 853;
Best Local Similarity 23.1%; Pred. No. 0.055; Indels 87; Gaps 18;
Matches 97; Conservative 80; Mismatches 156; Indels 87; Gaps 18;

QY 34 BESSWDSLSCLTWSQKDVCLPKAAHQKEIDKING-----KLEGSVPKDGILLKANCQMK 87
DB 57 DRSSPNSKSVERRSPK-LPTPEKSQARVAVKGTESQTTTLRSQIKEDLKANERIS 115
QY 88 VSIPTKALELMDQTFKAPPEKPSAFAPAEMQK-----SVPNKA 128
DB 116 SLEKDKAKALDELQAKKEAQVTLKLDLAKAQHVEENSEBIKQFAVEAGIEAVQNE 175
QY 129 LELKNE-OTLR-----ADEILPESKQDYEE--SSWDSLSCLTWSQ-KDVCLPKAT 177
DB 176 BELKKELETVKNQHASDAAVAVRQELKINEELAAAFDAKS--KALSQADASKTABI 233
QY 178 HKEIDKINGKL-----ESPNDGFLKAPCRMKVSIPTKALELMDQTFKAP 226
DB 234 HAQKVDILSELTRUKALDSTREKTAISDNEMWAKLEDEIVLKKDLB--SARGFEAEV 291
QY 227 PEKPSAFAPA-----IEMQKSVPNKALELKNEOTLRADQMPPESSKQKNVEENSWSDESLE 283
DB 292 KEKEMIVKLVNDLEAAKMAESNAHSLSNW-----QSKAKELEEQLEANKLER 341
QY 284 TVSQKDVCPKATHQKEMKIGSKLEDS--TSLSKILDTHSCERARELQKH---CEQCT 339
DB 342 SAS-----VSLESVMKQSGNSDKLHDTETITDLKERIVTLTETTVAKQKEDLEVSEQR 396
QY 340 GRMEQMKKFCVCLKKLSAKETKSOLE-----NOKVWEOELCS-VRLFTLMKMKIIS 392
DB 397 GSVEE-----EVSKEKEVEKLKSELTETVKEKNRALKAEQDATSRVORLSEKSKLIS 450

RESULT 10
S49461
synaptonemal complex protein 1 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S49461; S59599
R:Julien, S.; Luc, M.; Francols, C.
submitted to the EMBL Data Library, October 1994
A:Description: Cloning and sequencing of the murine SCP1 cDNA.
A:Reference number: S49461
A:Accession: S49461
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-993 <JUL>
A:Cross-references: UNIPROT:Q62209; EMBL:Z38118; NID:gl360015; PIDN:CAA86262.1; PID:g5586
R:Sage, J.; Martin, L.; Cuzin, F.; Rassoulzadegan, M.
Biochim. Biophys. Acta 1263, 258-260, 1995
A:Title: cDNA sequence of the murine synaptonemal complex protein 1 (SCP1).
A:Reference number: S59599; MUID:96004899; PMID:7548215
A:Accession: S59599
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-993 <SAG>
A:Cross-references: EMBL:Z38118; NID:gl360015; PIDN:CAA86262.1; PID:g558603

Query Match 8.0%; Score 164.5; DB 2; Length 993;
Best Local Similarity 21.3%; Pred. No. 0.075;
Matches 93; Conservative 90; Mismatches 146; Indels 107; Gaps 20;

QY 1 MOKSVPNKALELKN-EQTLRADEILPSESQK-----KDYESSWDSSESLCETVS----- 48
DB 419 MTKFNKKEVELEBELKNLAEQDKLLDEKKQVEKLAELQEQELTFLLETREREVHDL 478
QY 49 QKDVCLPKAAHQ---KEIDKINGKLEGSVPKDGILLKANCQMKVSIPTKALELMDQTFKA 105
DB 479 QEQTVTVTTSQHYLKQVEEMKTELEKEKLNTELTAASCDML-----LLENKKFVQ 529
QY 106 EPPEKPSAFAPAEMQK-----SVPNKALELKN-EQTLRADEILP 144
DB 530 E-----ASDMALELKHQEDIIINCKQEFERLLKQIENLEEKEMHLRDELSVRKEFTQ 583
QY 145 SESQKQDYESSWDSSESL-CETV---SQKDVCLPKATH-QKEIDKINGKLESPNDGFL 199
DB 584 GDEVCKLDKSEENARSIECEVLKKEKQMKILESCKNLKQVENKSNIEELHQENKTL 643
QY 200 KAPCRMKV-----SIPTKALELMDQTFKAPPEKPSAFAPAEMQ----- 240
DB 644 KKSASAEIKQLNAYEIKVSKLEL-ELESTKQRFEEWNTNNYQKEIENKKISEGKLLGEVEK 702
QY 241 -KSVPNKALELKNQOTLRAD---QMPPESSKQKN-----VEENSWSDESLSRETVSQDV 290
DB 703 AKATVDEAVKLQKEIDRLCQHKIAEMVALMEKHQYDKIVEER--DSEL----- 750
QY 291 CVPKATHQKEMKIGSKLEDSLSKILDTHSCERARELQKHCEQCTGKMEQMKKFC 350
DB 751 ---GLYKREOQSSAKIALETLSNIRNELVSLKQKEIEKEBE---KUMAKENTA 803
QY 351 VLKKKLSAKETKSOL 366
DB 804 ILKDK--KDKKIQASL 817

RESULT 11
T13030
microtubule binding protein D-CLIP-190 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13030
R:Lantz, V.A.; Miller, K.G.
J. Cell Biol. 140, 897-910, 1998
A:Title: A class VI unconventional myosin is associated with a homologue of a microtubule
A:Reference number: Z17588; MUID:98139549; PMID:9472041
A:Accession: T13030
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1690 <LAN>
A:Cross-references: EMBL:AF041382; NID:g2773362; PID:g2773363; PIDN:AAB96783.1
A:Experimental source: strain Oregon R
C:Genetics:
A:Cross-references: FlyBase:FBgn0020503
C:Keywords: cytoskeleton

Query Match 8.0%; Score 164.5; DB 2; Length 1690;
Best Local Similarity 20.9%; Pred. No. 0.14;
Matches 84; Conservative 75; Mismatches 149; Indels 93; Gaps 13;

QY 5 VPKNALELKNQOTLRADILPSESQKQDY-----EESWDSSESLCETVS-----QKDV 53
DB 1122 VANANISATNAELSTVLEVLQAEKSTNHFELFEMEADMNSERLIEKVTGIKEELKETH 1181
QY 54 LPKAAHQKEIDKINGKLEGSVPKDGILLKANCQMK-----VSIPTKALELMDQTFKA 106
DB 1182 LQDERQKKEFELEELKQAOQSEKQLQESQTSKEKLEIQQLQELQDSVQKQELVQ 1241
QY 107 -PPEKPSAFAPAEMQKSVPNKA-LELKNQOTL---RADEILPSESQKQDYESSWDS 161
DB 1242 NLEEKVRESSIIEAQNTKLNSNVQLENKTSCLKETQDLLESQKKEQLQE----- 1294

```

Qy 162 LCETVSQKDVCCLKPATHQEKIDKINGKLESPPNDGFLKAPCRMKVSIPTKALELMDMQT 221
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1295 -----EAAKLSGELQVQEQANGDIK-DSLVKVKBELVKVLE----- 1328
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 222 FKAEPPEKPSAFEPAIEMQKSVFNKALEL-----KNEOTLRADQMPSESKQKNVENSW 276
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1329 -----EKLRQATSQLDAQATNKELQELLVKSGENGLQGESLAVTEKQLOQLQANG- 1381
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 277 DSESLRETVSQKQVCPKATHQKEMDKISGLKEDSTLSKILDTIHSCEARELQKDHCE 336
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1382 ---ELKEALCQK-----ENGLKELQKLDSENTV-----LESQKKSHN 1416
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy 337 QCTGKQKQEMKKFCVLKKKLSAEKIKSOLENKQKVKEQEL 377
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1417 EIQDKLEAQQAQKERTLOEETSKLAELQSLQKQANLELQKSL 1457
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 12
F84730
Probable myosin heavy chain [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: F84730
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: F84730
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1269 <STO>
A/Cross-references: GB:AE002093; NID:g6598483; PIDN:AAC69932.2; GSPDB:GN00139
C/Genetics:
A/Gene: At2g32240
A/Map position: 2

```

A26655
myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N;Contains: myosin ATPase (EC 3.6.4.1)
C;Species: Dictyostelium discoideum
C;Date: 05-Oct-1988 #sequence revision 05-Oct-1988 #text_change 09-Jul-2004
C;Accession: A26655; A24728; S00250
R;Warrick, H.M.; De Lozanne, A.; Levinand, L.A.; Spudich, J.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A;Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A;Reference number: A26655; MUID:87092266; PMID:3540939
A;Accession: A26655
A;Molecule type: DNA
A;Residues: 1-2116 <WAR>
A;Cross-references: UNIPROT:P08799; GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1; F:
P;DeLozanne, A.; Lewis, M.; Spudich, J.A.; Levinand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A;Reference number: A24728; MUID:86016788; PMID:3901008
A;Accession: A24728
A;Molecule type: mRNA
A;Residues: 2035-2116
R;Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A;Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium
A;Reference number: S00250; MUID:86112226; PMID:2828113
A;Accession: S00250
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1734-1893 <WAG>
C;Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phospho
F;1-818/Domain: globular head <HED>
F;89-747/Domain: myosin motor domain homology <MMOT>
F;179-186/Region: nucleotide-binding motif A (P-loop)
F;819-2116/Domain: alpha-helical rod <ROD>

RESULT 13

RESULT 14

I49464
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49464; I49463; I49462; I49461; I49604
R:Quinn-Laquer, B.K.; Kennedy, J.E.; Wei, S.J.; Beisel, K.W.
Genomics 13, 176-188, 1992
A:Title: Characterization of the allelic differences in the mouse cardiac alpha-myosin h
A:Reference number: A38207; MUID:92250040; PMID:1577481
A:Accession: I49464
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1938 <RES>
A:Cross-references: UNIPROT:Q02566; GB:M76601; NID:g191623; PIDN:AAA37162.1; PID:g191624
A:Accession: I49463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1938 <RES>
A:Cross-references: GB:M76600; NID:g191621; PIDN:AAA37161.1; PID:g191622
A:Accession: I49462
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-544, 'A', 546-1938 <RE4>
A:Cross-references: GB:M76598; NID:g191617; PIDN:AAA37159.1; PID:g191618
R:Gulick, J.; Subramaniam, A.; Neumann, J.; Robbins, J.
J. Biol. Chem. 266, 9180-9185, 1991
A:Title: Isolation and characterization of the mouse cardiac myosin heavy chain genes.
A:Reference number: I49604; MUID:91225025; PMID:2026617
A:Accession: I49604
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-67 <RE5>
A:Cross-references: GB:M62404; NID:g192609; PIDN:AAA37424.1; PID:g192610
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: ATP; cardiac muscle; heart; nucleotide binding; P-loop
F:88-768/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)
Query Match 7.9%; Score 162.5; DB 2; Length 1938;
Best Local Similarity 24.0%; Pred. No. 0.21;
Matches 101; Conservative 79; Mismatches 164; Indels 77; Gaps 22;
QY 2 QKSVNKALELKNQTLRADEILPSESQKQYESSWDSSELCETVSQKQVCLPKAAHQK 61
DB 848 EKEMAN---MKEEGRVVDKALESEARKELEEK---MVSLQ--EKNDLQVQAEQ- 897
QY 62 EIDKINGKLEGSFVKDGLLKANGCMKVSIPTKALELMDMTFKAPPEKPSAFE-PAIEM 120
DB 898 --DNLDAEERC---DQLIKNKIQLEAKVEMTERLEDEENNAELTAKRKLEDECSL 952
QY 121 QKSVNKALELKNQTLRADEILPSESQKQYESSWDSSELCETVSQKQVCLPKATHQ 179
DB 953 KXIDIDLETLTAKVEKHA-----TENKVNLTTEEMAGLDEIATKIKKALQEA-HQ 1006
QY 180 K-----EIDKING-----KLEESPDN-DGFL--KAPCRMKVSIPTKALE----- 215
DB 1007 QALDDLOAEEDKNTLTSKVKLEQQVDLEGSLEQKVKMDLERAKKLEGLDKLTQE 1066
QY 216 -LMDMTFKAPPEKPSAFEPAIEMQKSVNKALELKNQTLRADQMFPSQKQVNEEN 274
DB 1067 SIMDLENDKLOLEKLEKKKEFDISQNS-----KIEDQAL-ALQL-----QKKLKEN 1113
QY 275 SWDSSELRVTSQKQVCPVPAH-----QKMDKISGLKEDSTSLSKILDTIHCERAE 329
DB 1114 QARIEELEELEARTARAKVKLRSDLSRELESEERLEAGGATSVQIEMNK-KREAE 1172
QY 330 LQK--DHCEQCTGME-----QMKK---FCVLKKLSEAKETKSLQENQKVKWEQLCS 379

DB 1173 FQWRRLDEATLQHEATAAALRKKHADSVAEELGQIDNLRVQKLEKESEFKLELDD 1232
QY 380 V 380
DB 1233 V 1233
RESULT 15
T18278
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18278; T30579
R:Hammer III, J.A.; Jung, G.
J. Biol. Chem. 271, 7120-7127, 1996
A:Title: The sequence of the dictyostelium myo J heavy chain gene predicts a novel, dimer
A:Reference number: Z18854; MUID:96215148; PMID:8636147
A:Accession: T18278
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2245 <HAM>
A:Cross-references: UNIPROT:P54697; EMBL:U42409; NID:g1150765; PID:g1150766; PIDN:AAA8511
R:Titus, M.A.; Kuapa, A.; Loomis, W.F.
Proc. Natl. Acad. Sci. U.S.A. 91, 9446-9450, 1994
A:Title: Discovery of myosin genes by physical mapping in Dictyostelium.
A:Reference number: Z20873; MUID:95023928; PMID:7937787
A:Accession: T30579
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-190, 'F', 192-283, 'T', 285-290, 'R', 292-331, 'IE', 'WM', 337-338, 'LK', 342, 'YRMSK',
A:Cross-references: EMBL:L35322; NID:g1039360; PID:g1039361; PIDN:AAA79858.1
C:Genetics:
A:Gene: myoJ
A:Introns: 257/3; 307/3
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: ATP; P-loop
F:84-809/Domain: myosin motor domain homology <MMO>
Query Match 7.9%; Score 162.5; DB 2; Length 2245;
Best Local Similarity 21.8%; Pred. No. 0.25;
Matches 98; Conservative 73; Mismatches 166; Indels 113; Gaps 18;
QY 7 NKALELKNQTLRADEILPS-----ESKQYESSWDS-SELCETVSQKQVCLPKAAHQK 57
DB 1071 NKKLEKLDSDQHDSTEKLQSQFNETEQQLQKQKQSELSKLSKTTQQLDFNKQEFDR 1130
QY 58 AHQKEIDKINGKLEGSFVKDGLLKAN-----CGMKVSIPTKALELMDMTFKAP 107
DB 1131 SQERDTNTNNQLEIQQLK-----KANSTLEEDYFSLGIRDNLQVLELRDENQLIKER 1186
QY 108 ----PEKPSAFEPAIEMQKSVNKALELKNQTLRADEILPSESQKQ-----DYESSW 158
DB 1187 LDSLQQQSQSQFQSGAALKEQLQVQSEQLIKLSEKLGSEBAKQINQLELELTD 1246
QY 159 SEELCETVSQKQVCLPKATHQKEIDKINGKLEESPDNGFLKAPC-RMKVS----- 208
DB 1247 HSKSLQIQQLQ-----LTEQSNK-INKLKGLEYQDEKKQLQQLLEIRIKQSQSVEDEKN 1301
QY 209 -----IPTKALELMDMTFKAPPEKPSAFEPAI-EMQKSVNKALELKNQTLRADQMF 262
DB 1302 SLITQLTTVKFESTQVSTNVSHQKEKITTLKSTIBELNKSIGKLAQKQNKD----- 1353
QY 263 PSESQKQVNEENSWDSSELRVTSQKQVCPVPAHQMCKDKISGLKEDSTSLSKILDTI 322
DB 1354 -DEIRKIQFELNDKQKQFTRQTKPEFSDL-----SQSQSIDRPKSEI-----TIN 1396
QY 323 SCERARELQKHCE-----QCTGKMEOMKKKFCVLKK 354
DB 1397 SLERNETLKSDFERVQSQSLKQQRCCQYKDTINRLENEVKQLTQLKERNFENFPAKE 1456
QY 355 KLSE-----AKETKSLQENQKVKWEQL 377

Db 1457 QNSNQTESVYLKEVTTQMQQNSRIEREL 1486

Search completed: August 1, 2005, 22:23:13
Job time : 26.5213 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:23:28 ; Search time 587.051 Seconds
(without alignments)
3196.585 Million cell updates/sec

Title: US-09-489-079-25

Perfect score: 1632

Sequence: 1 MGTRALQCEVSHHENY.....NHLKNRIYQEKKAETENS 317

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q/cn2_1/USFTO_spool_p/US09489079/runat_01082005_141249_13290/app_query.fasta_1.1635
-DB=N Geneseq 16Dec04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1.1 1418 @runat_01082005_141249_13290 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1632	100.0	1665	4 AAF17692	Aaf17692 Human bre
2	1632	100.0	1665	4 AAI67216	Aai67216 B726P sec
3	1632	100.0	1665	4 AAS47122	Aas47122 Human cDN
4	1632	100.0	1665	6 ABT08777	Abt08777 Human bre
5	1632	100.0	1665	6 ABS63723	Abs63723 Human bre

6	1632	100.0	1665	10 ABT32935	Abt32935 Human tum
7	1632	100.0	1665	11 ADL92840	Adl92840 Human bre
8	1632	100.0	1665	12 ADE44130	Ade44130 Human cDN
9	1632	100.0	2043	9 ADB83986	Adb83986 Human bre
10	1632	100.0	2307	4 AAF17980	Aaf17980 Human bre
11	1632	100.0	2307	4 AAI67223	Aai67223 B726P spl
12	1632	100.0	2307	4 AAS47410	Aas47410 Human cDN
13	1632	100.0	2307	6 ABS64011	Abs64011 Human bre
14	1632	100.0	2307	10 ABT33223	Abt33223 Human tum
15	1632	100.0	2307	11 ADL93130	Adl93130 Human bre
16	1632	100.0	2307	12 ADE44420	Ade44420 Human cDN
17	1632	100.0	2683	3 AAS59015	Aas59015 Nucleotid
18	1632	100.0	2683	3 AAS59014	Aas59014 Nucleotid
19	1632	100.0	3681	4 AAF17975	Aaf17975 Human bre
20	1632	100.0	3681	4 AAI67218	Aai67218 B726P spl
21	1632	100.0	3681	4 AAS47405	Aas47405 Human cDN
22	1632	100.0	3681	6 ABS64006	Abs64006 Human bre
23	1632	100.0	3681	10 ABT33218	Abt33218 Human tum
24	1632	100.0	3681	11 ADL93125	Adl93125 Human bre
25	1632	100.0	3681	12 ADE44415	Ade44415 Human cDN
26	1622	99.4	972	4 AAS47417	Aas47417 Human cDN
27	1622	99.4	972	6 ABS64018	Abs64018 Human bre
28	1622	99.4	972	10 ABT33230	Abt33230 Human tum
29	1622	99.4	972	11 ADL93144	Adl93144 Human bre
30	1622	99.4	1206	4 AAF17690	Aaf17690 Human bre
31	1622	99.4	1206	4 AAI67215	Aai67215 B726P fir
32	1622	99.4	1206	4 AAS47120	Aas47120 Human cDN
33	1622	99.4	1206	6 ABT08775	Abt08775 Human bre
34	1622	99.4	1206	6 ABS63721	Abs63721 Human bre
35	1622	99.4	1206	10 ABT32933	Abt32933 Human tum
36	1622	99.4	1206	11 ADL92837	Adl92837 Human bre
37	1622	99.4	1206	12 ADE44127	Ade44127 Human cDN
38	1622	99.4	1233	4 AAS47423	Aas47423 Human bre
39	1622	99.4	1233	6 ABS64024	Abs64024 Human bre
40	1622	99.4	1233	10 ABT33236	Abt33236 Human tum
41	1622	99.4	1233	11 ADL93154	Adl93154 Human bre
42	1617	99.1	978	10 ABT33260	Abt33260 Human tum
43	1617	99.1	978	11 ADL93212	Adl93212 Human bre
44	1586	97.2	1681	4 AAF17693	Aaf17693 Human bre
45	1586	97.2	1681	4 AAI67217	Aai67217 B726P thi

ALIGNMENTS

RESULT 1
AAFI7692
ID AAF17692 standard; cDNA; 1665 BP.

AC AAF17692;

DT 13-MAR-2001 (first entry)

XX Human breast cancer associated B726P-74 coding sequence.

DE Human breast cancer associated gene; vaccine; diagnosis; therapy; ss.

XX Homo sapiens.

OS WO200060076-A2.

PN 12-OCT-2000.

PD 15-FEB-2000; 2000WO-US005308.

XX 02-APR-1999; 99US-00285480.

XX 23-JUN-1999; 99US-00339338.

XX 02-SEP-1999; 99US-00389681.

XX 03-NOV-1999; 99US-00433826.

XX (CORI-) CORIXA CORP.

XX Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;

XX

DR WPI; 2001-122627/13.
DR P-PSDB; AAB50243.
XX An isolated polypeptide useful for the treatment and diagnosis of tumors
PT e.g. breast cancer comprises at least an immunogenic portion of a breast
PT tumor protein.
XX
XX
PS Claim 6; Page 143; 238pp; English.
XX
CC The present invention provides the coding sequences and some protein
CC sequences of proteins associated with breast cancer in humans. These
CC sequences can be used in the diagnosis and treatment of cancers,
CC particularly breast tumors
XX
SQ Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.1e-136 Length: 1665
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-25 (1-317) x AAF17692 (1-1665)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 459 ATGGGAACAAGAGCTGCTGCAGTGTGAGGTTCTCACACTCATGAAATGAAATATATCTC 518

Qy 21 LeuHisGluAsnCysMetLeuLeuLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
Db 519 TTACATGAAATTCGATGTTGAAAAGGAATTCGCATGCTTAAACTGGAATAGCCACA 578

Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db 579 CTGAAACACCAATACCAGAAAGGAAATTAATACCTTGAGGACATTAAGATTAAAA 638

Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluSerLeuThrLysArg 80
Db 639 GAAAGAATGCTGAACTTCAGATCACCTTAAACTGAAAGAGGAATCATTAACATAAAGG 698

Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db 699 GCATCTCATATATAGTGGCAGCTTAAAGTCTTGATAGCTGAGAACACATGCTCACTCT 758

Qy 101 LysLeuLysGluLysGlnAspLysGluLeuLeuGluAlaGluIleGluSerHisGluPro 120
Db 759 AAATTGAAGGAAAAACAAGACAAAGAAATACCTAGAGGCGAGAAATTGAATCACACATCCT 818

Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTCTGCTGTGTACAGACCATGATCAAAATGTCATCAAGAAAAAGTCAAGAA 878

Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 879 CTGCTTCCACATGTCAGGAGATGCTGTTTGCAGAAATAATGAAATGTTGATGAGT 938

Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTACGATATATAACAATGAGTGTCTCCATCAACCATCTTCTGAAGCTCAAAAGGAATCC 998

Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 999 AAAAGCCTTAAATAATCTCAATATATGCCGAGATGCTCTAAGAGAAAAATACATGTTGT 1058

Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGAACATGTCACAAAGAGACCAAGTGAACACAGTGTCAATGAAGGAGGCTGAACAC 1118

Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db 1119 ATGTATCAAAACGAACAAGATAATGTGAAACAAACACACTGAACAGCAGGAGTCTTAGAT 1178

Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
Db 1179 CAGAAATATTTCACTACAAAGCAAAATATGTGGCTTCAACAGCAATTAGTTCAATGCA 1238

Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1239 CATAAGAAAGCTGACACAAACAAAGCAAGATAACAATTCATATTCATTTCTTGAGAGGAAA 1298

Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluLulePheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTCTTAAAGAGAAATGAGGAGATATTTAATTAACAATCAATTTA 1358

Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1359 AAAAACCGTATATATCAATATGAAAAAGAAAGACAGAAACAGAAACTCA 1409

RESULT 2
AAI67216
ID AAI67216 standard; cDNA; 1665 BP.
XX
AC AAI67216;
XX
DT 11-FEB-2002 (first entry)
XX
DE B726P second splice variant encoding cDNA.
XX
KW Genetic subtraction; DNA microarray analysis; polymerase chain reaction;
KW cancer; B726P; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..540
FT /*tag= a
FT /product= "B726P splice variant"
XX
PN WO200175171-A2.
XX
PD 11-OCT-2001.
XX
PF 02-APR-2001; 2001WO-US010631.
XX
PR 03-APR-2000; 2000US-0194241P.
PR 20-JUL-2000; 2000US-0219862P.
PR 27-JUL-2000; 2000US-0221300P.
PR 18-DEC-2000; 2000US-0256592P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Dillon DC, Molesh DA, Xu J, Zehentner B, Persing DH;
XX WPI; 2001-626449/72.
DR P-PSDB; AAG65981.
XX
PT Identifying tissue (tumor)-specific polynucleotides overexpressed in
PT tissue of interest as compared to control tissue, for detecting cancer
PT cells in patient, comprises DNA microarray analysis or quantitative
PT polymerase chain reaction.
XX
PS Claim 4; Page 101-102; 127pp; English.
XX
CC The invention relates to identifying tissue-specific polynucleotides (P)
CC that involves performing a genetic subtraction to identify pool of (P)
CC from tissue of interest (TI), performing DNA microarray analysis to
CC identify first subset of polynucleotides (SP1) at least 2-fold over
CC expressed in TI, and performing quantitative polymerase chain reaction
CC (PCR) analysis on SP1 to identify second subset of (P). The method is
CC useful for determining the presence or absence of a cancer cell in a
CC patient, monitoring the progression of cancer in a patient using a
CC biological sample such as blood, serum, lymph nodes, bone marrow, sputum,
CC urine or a tumour biopsy sample. The methods are useful for determining
CC the presence or absence of or monitoring progression of prostate, breast,

CC colon, ovarian, lung, head and neck, lymphoma, leukemia, melanoma, liver,
 CC gastric, kidney, bladder, pancreatic or endometrial cancer. The present
 CC sequence represents a cDNA encoding a B726P second splice variant
 XX
 SQ Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2.1e-136 Length: 1665
 Score: 1632.00 Matches: 317
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-25 (1-317) x AAI67216 (1-1665)

QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
 DB 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 518
 QY 21 LeuHisGluAsnCysMetLeuLysLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
 DB 519 TTACATGAATATGCAATGTTGAAAAGGAATGTCATGCTAAACTGGAATAGCCACA 578
 QY 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspLysLeuLys 60
 DB 579 CTGAAACACCAATACCAGGAAAAGGAAAATAAATACTTTGAGGACATTAAGATTTTAAA 638
 QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
 DB 639 GAAAAGATGCTGAACTTCAGATGACCCCTAAAACCTGAAAGAGGAATCATTAATAAGG 698
 QY 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
 DB 699 GCATCTCAATATAGTGGGAGCTTAAAGTTCATAGCTGAGACACAATGCTCACTTCT 758
 QY 101 LysLeuLysGluLysGlnAspLysGluLeuLeuGluAlaGluLeuGluSerHisPro 120
 DB 759 AAATTGAAGAAAACACAGCAAGAAATACTAGAGCGAGAAATTAATCAACACATCTC 818
 QY 121 ArgLeuLysSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
 DB 819 AGACTGGCTTCTGCTGTACAGACCATGATCAAAATGTGACATCAAGAAAAGTCAAGAA 878
 QY 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
 DB 879 CTGCTTTTCCATTCAGAGAGATGCTGTTTGGCAAGAAAATGATGTTGATGTGAGT 938
 QY 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
 DB 939 AGTAGCATATATAACCAATGAGGTGCTCCATCAACCACTTTCTGAGCTCAAGGAAATCC 998
 QY 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
 DB 999 AAAAGCTTAAAAATTAATCTCAATATATGCGGAGATGCTCTAAGAGAAAATACATTGGT 1058
 QY 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
 DB 1059 TCAGAAATGTCACAAAGAGACCAACCTGGAACACACAGTGTCAATGAAAGAGCTGAAAC 1118
 QY 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
 DB 1119 ATGTATCAAAACGAACAGATATGTGAACAAACACACTGACAGCAGGAGTCTCTAGAT 1178
 QY 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
 DB 1179 CAGAAATATTTCAACTACAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1238
 QY 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
 DB 1239 CATAGAAAGCTGACAAACAAAAGCAGATAACAATGATATTTCTTGTAGAGAAA 1298
 QY 281 MetGlnHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300

DB 1299 ATGCAACATCATCTCTCTAAAGAGAAAATGAGGAGATATTAAATACAAATCACTTTA 1358
 QY 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
 DB 1359 AAAAACCGTATATATATATGTAATGAAAAGAGAAAGCAGAAACAGAAACTCA 1409
 RESULT 3
 ID AAS47122 standard; cDNA; 1665 BP.
 AC AAS47122;
 DT 18-DEC-2001 (first entry)
 DE Human cDNA clone B726P-74 encoding a breast cancer protein.
 KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
 KW gene therapy.
 OS Homo sapiens.
 PN WO200179286-A2.
 PD 25-OCT-2001.
 PF 12-APR-2001; 2001WO-US012164.
 PR 17-APR-2000; 2000US-00551621.
 PR 08-JUN-2000; 2000US-00590751.
 PR 22-JUN-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 PA (CORI-) CORIXA CORP.
 PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 DR WPI; 2001-611721/70.
 DR P-PSDB; AAU33344.
 XX Breast Tumor Proteins and nucleic acids useful for the prevention,
 diagnosis and treatment of breast cancer.
 PS Claim 1; Page 197-198; 297pp; English.
 CC The invention relates to isolated breast tumour proteins and nucleic
 acids that encode them, including immunogenic fragments of the proteins.
 CC Also included are expression vectors expressing the proteins, transformed
 cells and antibodies raised against the proteins or an antigen presenting
 cell expressing the protein. The proteins and nucleic acids may be used
 in the prevention, diagnosis and treatment of diseases associated with
 inappropriate breast tumour protein expression, i.e. breast tumours and
 breast cancer e.g. by gene therapy. The nucleic acids and their
 complements may also be used as DNA probes in diagnostic assays to detect
 and quantitate the presence of similar nucleic acids in samples, and
 therefore which patients may be in need of restorative therapy. The
 CC modulators (e.g. antagonists) of breast tumour protein expression and
 activity. The antibodies and antagonists may also be used to down
 regulate expression and activity. The antibodies may also be used as
 CC diagnostic agents for detecting the presence of the proteins in samples
 (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
 CC purification diagnostic techniques. The present sequence is a cDNA from a
 CC breast tumour cDNA library isolated by subtractive hybridisation against
 CC a normal breast cDNA library and encodes a breast tumour protein of the
 CC invention
 SQ Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2.1e-136 Length: 1665
 Score: 1632.00 Matches: 317
 Percent Similarity: 100.00% Conservativity: 0

Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 100.00%		Indels: 0
DB: 4		Gaps: 0
US-09-489-079-25 (1-317) x AAS47122 (1-1665)		
Qy	1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20	
Db	459 ATGGGAACAGAGCTCTGCAGTGTGAGGTTCTCACACTCATGAATGAATTAATCTC 518	
Qy	21 LeuHisGluAsnCysMetLeuLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40	
Db	519 TTATCATGAAATTCATGTTGAAAGAGAAATGCCATGCTAAATCTGGAATAGCCACA 578	
Qy	41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspLysLeuLys 60	
Db	579 CTGAACACCAATACAGGAAGAAATTAATACTTTGAGGACATTAAGATTTTAAAA 638	
Qy	61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysSerLeuThrLysArg 80	
Db	639 GAAAGAATGCTGAACCTTCAGATGACCTTAACCTGAAGAGGAATCATTAACCTAAAGG 698	
Qy	81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100	
Db	699 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGACACAATGCTCCTCTCT 758	
Qy	101 LysLeuLysGluLysGlnAspLysGluLeuGluLeuGluLeuGluSerHisHisPro 120	
Db	759 AAATTGAAGGAAAAACAAGACAAGAAATATCTAGAGGCAGAAATGGAATCACACCATCCT 818	
Qy	121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140	
Db	819 AGACTGCTCTGCTGTACAGACCATGATCAATTTGTGACATCAAGAAAAGTCAAGAA 878	
Qy	141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160	
Db	879 CCTGCTTTCACATTCGAGGAGATGCTTGTGTGCAAGAAAAATGAATGTGATGTGAGT 938	
Qy	161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180	
Db	939 AGTACGATATATACAAATGAGGTGCTCCATCAACCACTTCTGAAGCTCAAGAGAAATCC 998	
Qy	181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200	
Db	999 AAAAGCCTAAAAATTAATCTCAATTATGCGGAGATGCTCTAAGAGAAAATACATCTGTT 1058	
Qy	201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220	
Db	1059 TCAGAACATGCACAAAGAGACCAACGCGAAACACAGTGTCAATGAAGGAAGCTGAACAC 1118	
Qy	221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240	
Db	1119 ATGTATCAAAACGAAACGAATAATGTGAACAAACACACTGAACAGCAGGAGTCTCTAGAT 1178	
Qy	241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnLeuValHisAla 260	
Db	1179 CAGAAATTAATTTCAACTCAAGCAAAATATGTGGCTTCAACAGCAATAGTTCATGCA 1238	
Qy	261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspLysHisPheLeuGluArgLys 280	
Db	1239 CATAGAAGAGCTGACAAACAAAGCAAGATAACAATGATATTCATTTCTTGTAGAGGAAA 1298	
Qy	281 MetGlnHisIleLeuLysGluLysGluLysGluLysGluLysPheAsnTyrAsnAsnHisLeu 300	
Db	1299 ATGCAACATCATCTCTCTAAAGAGAAATAGAGAGATATTTAATTACAATAACCATTTA 1358	
Qy	301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317	
Db	1359 AAAAACCGTATATATCAATATGAAAAGAGAGAAACAGACAGAAACTCA 1409	
RESULT 4		
ABT08777		
ID ABT08777 standard; cDNA; 1665 BP.		
XX	ABT08777;	
AC	05-DEC-2002 (first entry)	
DT	Human breast antigen-related nucleotide #175.	
DE	Immunogenic protein; breast antigen; breast antigen-related DNA; vaccine;	
KW	breast cancer; adoptive immunotherapy; tumour-reactive T-cell; cancer;	
KW	ss.	
OS	Homo sapiens.	
XX	US2002102602-A1.	
PN	01-AUG-2002.	
PD	23-JUN-1999; 99US-00339338.	
PF	28-DEC-1998; 98US-00222575.	
XX	02-APR-1999; 99US-00285480.	
PR	(YUQI/) YUQIU J.	
PA	(DILL/) DILLON D C.	
PA	(MITC/) MITCHAM J L.	
PA	(XUJ/) XU J.	
XX	Yuqiu J, Dillon DC, Mitcham JL, Xu J;	
PI	WPI; 2002-697821/75.	
DR	P-PSDB; AAO15778.	
DR	Novel isolated polypeptide comprising an immunogenic portion of a breast	
PT	antigen, useful for treating breast cancer or for inhibiting the	
PT	development of breast cancer in a patient.	
XX	Claim 4; Page 78-79; 127pp; English.	
XX	The invention comprises a protein which contains an immunogenic portion	
CC	of a breast antigen. The invention also comprises breast antigen-related	
CC	DNA sequences. The protein of the invention is useful as a vaccine for	
CC	inhibiting the development of breast cancer. The protein of the invention	
CC	is used in adoptive immunotherapy for the treatment of cancer, and	
CC	generating/isolating tumour-reactive T-cells which can be administered to	
CC	a patient. The present DNA sequence represents a human breast antigen-	
CC	related nucleotide of the invention	
XX	Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;	
SQ		
Alignment Scores:		
Pred. No.:	2.1e-136	Length: 1665
Score:	1632.00	Matches: 317
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	6	Gaps: 0
US-09-489-079-25 (1-317) x ABT08777 (1-1665)		
Qy	1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20	
Db	459 ATGGGAACAGAGCTCTGCAGTGTGAGGTTCTCACACTCATGAATGAATTAATCTC 518	
Qy	21 LeuHisGluAsnCysMetLeuLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40	
Db	519 TTATCATGAAATTCATGTTGAAAGAGAAATTAATACTTTGAGGACATTAAGATTTTAAAA 578	
Qy	41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspLysLeuLys 60	
Db	579 CTGAACACCAATACAGGAAGAAATTAATACTTTGAGGACATTAAGATTTTAAAA 638	
Qy	61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysSerLeuThrLysArg 80	

Db 639 GAAAGAAATGCTGAACCTCAGATGACCTTAAACCTGAAAGAGGAATCAATTAACCTAAAGG 698
 Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
 Db 699 GCACTCTCAATATAGTGGCAGCTTAAAGTCTGTAGCTGAGACACACATGCTCACTTCT 758
 Qy 101 LysLeuLysGluLysGlnAspLysGluLeuLeuGluAlaGluLeuGluSerHisPro 120
 Db 759 AAATTGAAGGAAAAACAAGACAAAGAAATCTAGAGGCGAGAAATGAATCAACCATCTCT 818
 Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnLeuValThrSerArgLysSerGlnGlu 140
 Db 819 AGACTGGCTTCTGCTGTAAGACCATGATCAATGTAAGACATCAAGAAAGTCAAGAA 878
 Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
 Db 879 CTTGCTTTCCATTTGCGAGAGATCTGTTGTGCAAGAAATGAATGTTGATGTGAGT 938
 Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
 Db 939 AGTACGATATATAAATCAATGAGTGTCTCATCAACCATCTTCTGAAGCTCAAGGAAATCC 998
 Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
 Db 999 AAAGCCCTAAATAATTAATCTCAATATGCGGAGATGCTTAAAGAGAAATACATTTGGT 1058
 Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
 Db 1059 TCAGAAATGTCACAAAGAGACCAAGTGNAAACAGTGTCAATGAAGAGCTGAACAC 1118
 Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnSerLeuAsp 240
 Db 1119 ATGTATCAAAACGAACAGATAATGTGAACAAACACACTGAACAGCAGAGTCTCTAGAT 1178
 Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
 Db 1179 CAGAAATTAATTTCAACTCAAAAGCAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1238
 Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
 Db 1239 CATAGAAAGCTGACACAAAGAGCAGATTAACAATGATATTCATTTCTTGAGAGGAAA 1298
 Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluLeuPheAsnTyrAsnAsnHisLeu 300
 Db 1299 ATGCAACATCATCTCTTAAAGAGAAATGAGGAGATATTTAATTAACAATAACCATTTA 1358
 Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
 Db 1359 AAAAAACCGTATATATCAATATGAAAAAGAGAAAGCAGAAACAGAAACTCA 1409

RESULT 5
 ABS63723
 ID ABS63723 standard; cDNA; 1665 BP.
 XX
 AC ABS63723;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Human breast tumour polynucleotide #175.
 XX
 KW Human; breast tumour polynucleotide #175.
 KW vaccine.
 XX
 OS Homo sapiens.
 XX
 PN US2002085998-A1.
 XX
 PD 04-JUL-2002.
 XX
 PP 13-APR-2001; 2001US-00834759.
 XX
 XX 28-DEC-1998; 98US-00222575.
 PR 02-APR-1999; 99US-00285480.

PR 23-JUN-1999; 99US-00339338.
 PR 02-SEP-1999; 99US-00389681.
 PR 17-APR-2000; 99US-00433826.
 PR 08-JUN-2000; 2000US-00551621.
 PR 22-JUN-2000; 2000US-00590751.
 PR 20-JUL-2000; 2000US-00604287.
 PR 20-JUL-2000; 2000US-00620405.
 PA (CORI-) CORIXA CORP.
 XX
 XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
 PI Henderson RA;
 XX
 XX WPI; 2002-635657/68.
 DR P-PSDB; ABG78911.
 XX
 PT Novel breast cancer polynucleotides and polypeptides encoded by the
 PT polynucleotides, useful for detecting the presence of breast cancer in a
 PT patient, and in pharmaceutical compositions, for treating breast cancer.
 PS Claim 1; Page 107-108; 247pp; English.
 XX
 CC The invention relates to an isolated breast tumour polynucleotide and the
 CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
 CC detecting the presence of breast cancer in a patient, and in
 CC pharmaceutical compositions for treating breast cancer. The sequences are
 CC useful for stimulating an immune response in a patient and can therefore
 CC be used in production of vaccines. The sequences are also useful for
 CC detecting the presence of a cancer in a patient, by obtaining a
 CC biological sample from the patient, contacting the biological sample with
 CC a composition of the invention and detecting the amount of polynucleotide
 CC that hybridizes to the sample. This sequence represents a human breast
 CC tumour polynucleotide of the invention
 XX
 SQ Sequence 1665 BP; 650 A; 289 C; 311 G; 375 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2.1e-136 Length: 1665
 Score: 1632.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-489-079-25 (1-317) x ABS63723 (1-1665)
 Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
 Db 459 ATGGGAACACAGAGCTCTGCAGTGTGAGGTTCACACTCATGAAATGAAATATATCTC 518
 Qy 21 LeuHisGluAsnCysMetLeuLysLysGluLeuAlaMetLeuLysGluLeuAlaThr 40
 Db 519 TTACATGAAATTTGCATGTTGAAAAAGGAAATGCGCATGCTAAAACTGGAATAGCCACA 578
 Qy 41 LeuLysHisGlnTyrGlnGlnLysGluAsnLysTyrPheGluAspIleLysLeuLys 60
 Db 579 CTGAAACACCAATACCGAAGAAAGAAATAAATACTTTGAGGACATTAAGATTAAAAA 638
 Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
 Db 639 GAAAGAAATGCTGAACCTTCAAGTGCCTTAAACCTGAAAGAGGAATCATTAATAAAGG 698
 Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
 Db 699 GCATCTCAATATAGTGGCAGCTTAAAGTCTGTAGCTGAGACACACATGCTCACTTCT 758
 Qy 101 LysLeuLysGluLysGlnAspLysGluLeuGluAlaGluLeuGluSerHisPro 120
 Db 759 AAATTGAAGGAAAAACAAGACAAAGAAATCTAGAGGCGAGAAATGAATCAACCATCTCT 818
 Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnLeuValThrSerArgLysSerGlnGlu 140
 Db 819 AGACTGGCTTCTGCTGTAAGACCATGATCAATGTAAGACATCAAGAAAGTCAAGAA 878

141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
 879 CCTGCTTCCATGTCAGGAGATGCTGTTGTGCAAGAAATGAAATGTGATGAGT 938
 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
 939 AGTACGATATATACATGAGTGCTCCATCAACCACTTTCTGAAGCTCAAGGAATCC 998
 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
 999 AAAAGCCATAAAATTAATCTCAATTATGCCGAGATGCTCTAAGAGAAATACATTGGTT 1058
 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
 1059 TCAGAAATGTCACAAAGAGACCAACGCGTGAACACACAGTGTCAATGAAAGGAGCTGAACAC 1118
 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
 1119 ATGTATCNAACGACACAGATAATGTGNACAAACACACTGNACAGCAGGAGTCTCTAGAT 1178
 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
 1179 CAGAAATTAATTTCAACTACAAAGCAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1238
 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
 1239 CATAGAAAGCTGACAAACAAAGCAAGATAACAATGATATTCATTTCTTTGAGGAGAA 1298
 281 MetGlnHisLysLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
 1299 ATGCAACATCATCTCTTAAAGAGAAATGAGGAGATATTTAATTACAATAACCAITTA 1358
 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
 1359 AAAAACCGTATATATCAATATGATAAAGAGAGAAAGCAGAAACAGAAAACTCA 1409
 RESULT 6
 ID ABT32935 standard; DNA; 1665 BP.
 AC ABT32935;
 DT 15-MAY-2003 (first entry)
 XX Human tumour-related DNA sequence - SEQ ID No 178.
 DE Human; ds; vaccine; gene therapy; T cell stimulation; T cell expansion;
 KW tumour; breast cancer; cancer; immune response stimulation.
 XX Homo sapiens.
 OS WO200283956-A1.
 FN 24-OCT-2002.
 PD 15-APR-2002; 2002WO-US012378.
 PF 13-APR-2001; 2001US-00834759.
 PR 07-DEC-2001; 2001US-00007805.
 PR 13-FEB-2002; 2002US-00076622.
 XX (CORI-) CORIXA CORP.
 PA Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
 PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
 PI Vedvick TS, Mcneill PD, Durham M;
 XX WPI; 2003-103376/09.
 DR New polypeptide and polynucleotide useful for stimulating and/or
 XX expanding T cells specific for a tumor protein and treating breast
 PT cancer.

XX Example 1; Page 208-209; 375pp; English.
 CC The invention comprises a method of stimulating and/or expanding T cells
 CC specific for a tumour protein. The invention further comprises human
 CC nucleic acids and proteins that are associated with tumours (e.g. breast
 CC cancer). The method and sequences of the invention are useful for
 CC stimulating and/or expanding T cells specific for a tumour protein,
 CC detecting the presence of cancer, stimulating an immune response in a
 CC patient and treating breast cancer. The present DNA sequence represents a
 CC human tumour-related DNA sequence
 XX
 SQ Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;
 Alignment Scores: Length: 1665
 Pred. No.: 2.1e-136 Matches: 317
 Score: 1632.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 10
 US-09-489-079-25 (1-317) x ABT32935 (1-1665)
 Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
 Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATTAATCTC 518
 Qy 21 LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
 Db 519 TTACATGAAATTTGCATGTTGAAAAAGGAAATTCCTATGCTAAACCTGGAATAGCCACA 578
 Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
 Db 579 CTGAAACCAATACCAAGGAAAGGAAATATAATACCTTTGAGGACATTAAGATTTTAAAA 638
 Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLysLysGluGluSerLeuThrLysArg 80
 Db 639 GAAAAGAAATGCTGAACTTCAGATGACCTTAAACTGAAAGAGGAATCATTAACCTAAAAAG 698
 Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
 Db 699 GCATCTCATATATAGTGGCGAGCTTAAAGTTCTGATAGCTGAGAACAACAATGCTCACTTCT 758
 Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro 120
 Db 759 AAATTGAAGGAAAGAAACCAAGCAAGAAATACTAGAGCAGAAATTTGAATCACACCATCT 818
 Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
 Db 819 AGACTGGCTTCTGCTGTACAAAGACCATGATCAAAATGTTGACATCAAGAAAAAGTCAAGAA 878
 Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
 Db 879 CCTGCTTTCACATATGCAAGAGATGCTTGTGTTGCAAGAAAAATGAAATGTGATGTGAGT 938
 Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
 Db 939 AGTACGATATATACAAATGAGGTGCTCCATCAACCACTTTCTGAAGCTCAAGGAATTC 998
 Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
 Db 999 AAAAGCCATAAAATTAATCTCAATTATGCCGAGATGCTCTAAGAGAAATACATTGGTT 1058
 Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
 Db 1059 TCAGAAATGTCACAAAGAGACCAACGCGTGAACACACAGTGTCAATGAAAGGAGCTGAACAC 1118
 Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
 Db 1119 ATGTATCNAACGACACAGATAATGTGNACAAACACACTGNACAGCAGGAGTCTCTAGAT 1178
 Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260

```
Db 1179 CAGAAATATTTCACAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1239 CATAGAAGAGCTGACACAAAGCAAGATAACAATTGATATTTCATTTCTTGAGAGGAAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTCTAAAGAGAAAATGAGGAGATATTAAATTACAATAACCATTTA 1358
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1359 AAAAACCGTATATATCAATATCAAAAGAGAAAGAGAGAAACAGAAAACCTCA 1409

RESULT 7
ADL92840
ID ADL92840 standard; cDNA; 1665 BP.
XX
AC ADL92840;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human breast cancer-associated polypeptide cDNA #175.
XX
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human; ss;
KW gene.
XX
OS Homo sapiens.
XX
PN US2003166022-A1.
XX
PD 04-SEP-2003.
XX
PF 15-APR-2002; 2002US-00124805.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
PR 13-FEB-2002; 2002US-00076622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Houghton RL, Sleath PR, Persing DH;
XX
DR WPI; 2003-874918/81.
XX
DR P-PSDB; ADL92841.
XX
PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX
XX
XX Example 1; SEQ ID NO 178; 294pp; English.
XX
PS The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents cDNA encoding a human breast cancer-associated
CC polypeptide.
XX
SQ Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.1e-136 Length: 1665
Score: 1632.00 Matches: 317
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
US-09-489-079-25 (1-317) x ADL92840 (1-1665)
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrIleu 20
Db 459 ATGGGAACAAGAGCTCTGCAAGTGTGAGGTTTCTCACACTCATGAAAATGAAAATTTATCTC 518
Qy 21 LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysAlaLeuGluIleAlaThr 40
Db 519 TTACATGAAAATTGCATGTTGAAAAGGAAAATGCCCATGCTAAAACTGGAATAGCCACA 578
Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db 579 CTGAACACCCATATACAGGAAGAAAATAAATACTTTGAGGACATTAAGATTTTAAAA 638
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLysLysGluGluSerLeuThrLysArg 80
Db 639 GAAAAGAAATGCTGAACCTTCAGATGACCTTAAACTGAAAGAGGAATCAATTAACAAAAAG 698
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db 699 GCATCTCAATATATGTGGCAGCTTAAAGTTCTGATAGCTGAGACACAAATGCTCACTTCT 758
Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro 120
Db 759 AAATTGAAGAAAACAAGACAAAGAAATATCTAGAGGCAAGAAATTTGAATCACACCATCT 818
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTTCTGCTGTACAAAGACCATGATCAAAATTTGTACATCAAGAAAAAGTCAAGAA 878
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 879 CCTGCTTTCCACATTTGACAGGAGATGCTTTGTTGCAAGAAAAATGAATTTGATGTAGT 938
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTAGATATATACATATGAGTGCTCCATCAACCATCTTCTGAAGCTCAAGGAAATCC 998
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 999 AAAAGCCTAAAAATTAATCTCAATTTATCGCGAGATGCTCTAAGAGAAAAATACATTGGTT 1058
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGAACATGCAAAAGAGACCAACGTGTAACACACAGTGTCAATGAAGGAAAGCTGAACAC 1118
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db 1119 ATGTATCAAAACGAAACAAGATATGTGAACAACAACACTGAACAGCAGAGTCTCTAGAT 1178
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
Db 1179 CAGAAATTTATTTCACATCAAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1239 CATAGAAGAGCTGACACAAAGCAAGATAACAATTGATATTTCATTTCTTGAGAGGAAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTCTAAAGAGAAAATGAGGAGATATTAAATTACAATAACCATTTA 1358
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1359 AAAAACCGTATATATCAATATGAAAAGAGAAAGAGAGAAACAGAAAACCTCA 1409
RESULT 8
ADE44130
```

ID ADE44130 standard; cDNA; 1665 BP.
 AC ADE44130;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human cDNA associated with breast cancer #175.
 XX
 KW human; ss; gene; breast tumour; cancer; vaccine; T cell stimulator;
 KW T cell expander.
 XX
 XX Homo sapiens.
 XX
 XX US2003104366-A1.
 XX
 XX 05-JUN-2003.
 PD
 XX
 XX 17-APR-2000; 2000US-00551621.
 PF
 XX
 XX 28-DEC-1998; 98US-00222575.
 PR
 XX 02-APR-1999; 99US-00285480.
 PR
 XX 23-JUN-1999; 99US-00339338.
 PR
 XX 02-SEP-1999; 99US-00389681.
 PR
 XX 03-NOV-1999; 99US-00433826.
 XX
 XX (JIAN/) JIANG Y.
 PA (DILL/) DILLON D C.
 PA (MITC/) MITCHAM J L.
 PA (XUJJ/) XU J.
 PA (HARL/) HARLOCKER S L.
 XX
 XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
 PI
 XX WPI; 2004-020270/02.
 DR
 XX P-PSDB; ADE44131.
 DR
 XX
 PT Novel isolated polypeptide comprising immunogenic portion of breast tumor
 PT protein or its variant, useful for formulating vaccines for inhibiting
 PT cancer development in a patient.
 XX
 XX Claim 5; SEQ ID NO 178; 217bp; English.
 PS
 XX The invention relates to an isolated polypeptide comprising at least an
 CC immunogenic portion of a breast tumour protein. The polynucleotide, its
 CC polypeptide, its antibody, a pharmaceutical composition comprising the
 CC fusion protein or the polynucleotide encoding it, a vaccine comprising
 CC the fusion protein or the polynucleotide encoding it, an isolated T cell
 CC population comprising T cells specific for a breast tumour protein, and a
 CC method for removing tumour cells from a biological sample is useful for
 CC inhibiting the development of a cancer in a patient. The polypeptide is
 CC useful for stimulating and/or expanding T cells specific for a breast
 CC tumour protein. Stimulating and/or expanding T cells specific for a
 CC breast tumour protein is useful for inhibiting the development of a
 CC cancer in a patient. The method additionally involves the step of cloning
 CC at least one proliferated cell and then administering the cloned T cells
 CC to the patient. The present sequence represents a cDNA associated with
 CC breast cancer.
 XX
 XX SQ Sequence 1665 BP; 690 A; 289 C; 311 G; 375 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,1e-136 Length: 1665
 Score: 1632.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0
 US-09-489-079-25 (1-317) x ADE44130 (1-1665)
 Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAenGluAenTyrIeu 20
 Db 459 ATGGGAACAGAGCTCTGCACTGTGAGGTTCTCACACTCATGAAATGAAATATTATCTC 518

Qy 21 LeuHisGluAenCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
 Db 519 TTACATGAAATTCGATGTTGAAAGAAAGAAATTCCTATGCTAAACTGGAAATAGCCACA 578
 Qy 41 LeuLysHisGlnTyrGlnGlnLysGluAenLysTyrPheGluAspIleLysIleLeuLys 60
 Db 579 CTGAACACCAATACACAGGAAAGGAAATAAATACTTTGAGGACATTAAGATTTTAAAA 638
 Qy 61 GluLysAenAlaGluLeuGlnMetThrLeuLysLysGluLysLeuThrLysArg 80
 Db 639 GAAAAGAAATCTGAACCTTCAGATGACCTTAAAGTCTGAAAGAGGAATCATTAACCTAAAGG 698
 Qy 81 AlaserGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAenThrMetLeuThrSer 100
 Db 699 GCATCTCAATATAGTGGGAGCTTAAAGTCTGATAGCTGAGAACACAAATGCTCACTTCT 758
 Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro 120
 Db 759 AAATTGAAGGAAAAAACAAGACAAAGAAATACTAGAGGCAGAAATTTGAATCACACCATCCT 818
 Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
 Db 819 AGACTGGCTTCTGCTGTACAAAGCCATGATCAAAATGTGTGATCATCAAGAAAGTCAAGAA 878
 Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAenValAspValSer 160
 Db 879 CTGTCTTTCACATTCGAGGAGATGCTTGTTCGAAAGAAATGATGTTGATGAGT 938
 Qy 161 SerThrIleTyrAenAenGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
 Db 939 AGTACGATATATAACAATGAGTGTCTCCATCAACCACTTTCTGAAGCTCAAGGAAATCC 998
 Qy 181 LysSerLeuLysIleAenLeuAenTyrAlaGlyAspAlaLeuArgGluAenThrLeuVal 200
 Db 999 AAAAGCCTAAAAATTAATCTCAATTTATGCCGAGATGCTCTAAAGAGAAATACATTCGTT 1058
 Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
 Db 1059 TCAGACATGCACAAAGAGACCAACGTCGAACACACAGTGTCAATGAAGGAGGAGTGAACAC 1118
 Qy 221 MetTyrGlnAenGluGlnAspAenValAenLysHisThrGluGlnGlnSerLeuAsp 240
 Db 1119 ATGTATCAAAACGAACAAGATAATGTGAACAAACACACTGAACAGCAGGAGTCTCTAGAT 1178
 Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAenMetTyrLeuGlnGlnGlnLeuValHisAla 260
 Db 1179 CAGAAATATTTCACACTACAAAGCAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1238
 Qy 261 HisLysLysAlaAspAenLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
 Db 1239 CATAGAAGAGCTGACACAAAGACAGATTAACATTCATATTCATTTCTTGAGAGGAA 1298
 Qy 281 MetGlnHisIleLeuLysGluLysAenGlnGluIlePheAenTyrAenAenHisLeu 300
 Db 1299 ATGCAACATCATCTCTCTTAAAGAGAAATGAGGAGATATTTAATTACAATAACCAITTA 1358
 Qy 301 LysAenArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAenSer 317
 Db 1359 AAAAACCGTATATATCAATATGAAAGAGAAAGAGGAGAAAGAGAAAGAAACTCA 1409
 RESULT 9
 ADB83986
 ID ADB83986 standard; cDNA; 2043 BP.
 XX
 AC ADB83986;
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX Human breast cancer diagnostic marker cDNA Incyte 411152.
 DE
 XX ss; cancer; neurodegenerative disorder; human; breast cancer;
 KW diagnostic marker.


```

XX OS Homo sapiens.
XX FH Location/Qualifiers
XX FT 56..778
XX FT /*tag= a
XX FT /partial
XX FT /transl_except= (pos:269..271,aa:Gly)
XX FT /transl_except= (pos:278..280,aa:Val)
XX FT /transl_except= (pos:281..283,aa:Lys)
XX FT /transl_except= (pos:290..292,aa:Leu)
XX FT /transl_except= (pos:302..304,aa:Asn)
XX FT /transl_except= (pos:308..310,aa:Gly)
XX FT /note= "No stop codon given"
XX PN US2003104418-A1.
XX PD 05-JUN-2003.
XX PF 25-APR-2002; 2002US-00133757.
XX PR 27-APR-2001; 2001US-0287153P.
XX PA (ZHAN/) ZHANG C.
XX PA (MAHI/) MAHINI B.
XX PA (WALK/) WALKER M G.
XX PI Zhang C, Mahini B, Walker MG;
XX DR WPI; 2003-687833/65.
XX DR P-PSDB; ADB83983.
XX PT New combination of polynucleotides, useful for preparing a composition
XX PS for diagnosing or treating cancer or neurodegenerative disorders.
XX PS Claim 4; Page 19-20; 22pp; English.
XX CC The invention related to a combination of polynucleotides. The
XX CC combination of polynucleotides is useful for preparing a composition for
XX CC diagnosing or treating cancer or neurodegenerative disorders. The present
XX CC sequence represents the human breast cancer diagnostic marker cDNA Incyte
XX CC 411152.
XX SQ Sequence 2043 BP; 822 A; 360 C; 377 G; 484 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 27e-136 Length: 2043
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-489-079-25 (1-317) x ADB83986 (1-2043)
QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
DB 814 ATGGGACACAGAGCTCTGAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 873
QY 21 LeuHisGluAsnCysMetLeuLysLysGluileAlaMetLeuLysLeuLeuAlaThr 40
DB 874 TTACATGAAAAATTGCTGTTGAAAGAGGAAATTCCTATGCTGAAATAGCCACA 933
QY 41 LeuLysHisGlnTyrGlnGlnLysGluAsnLysTyrPheGluAsnLysLysLysLys 60
DB 934 CTGAAACACCAATACCCAGGAAAGGAAATATAATCTTTGAGGACATTAAGATTTTAAAA 993
QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
DB 994 GAAAGAAATGCTGAATCTCAGATGACCCCTAAATCTGAAAGAGGAATCATTAATAAAGG 1053
QY 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuileAlaGluAsnThrMetLeuThrSer 100

```

```

DB 1054 GCATCTCAATATAGTGGCAGCTTAAGTTCTGATAGCTGAGAACACAAATGCTCACTTCT 1113
QY 101 LysLeuLysGluLysGlnAspLysGluLeuGluAlaGluileGluSerHisHisPro 120
DB 1114 AAATTGAGGAAAAACAAGACAAAGNAATACTAGAGGCAGAAATTTGAATTCACACCATCT 1173
QY 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
DB 1174 AGACTGGCTTCTGCTGTACAGACCATGATCAATTTGTGACATCAAGAAAAAGTCAAGAA 1233
QY 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
DB 1234 CTGCTTTCCATTCAGAGGAGATGCTTTGTCAGAGAAAAATGCAATGTTGATGTGAGT 1293
QY 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
DB 1294 AGTAGCATATATAACAATGAGGTGCTCCATCAACCACTTTCTGAAGCTCAAGGAAATCC 1353
QY 181 LysSerLeuLysLysLeuAsnTyrAlaGlyAspAlaLeuAtgGluAsnThrLeuVal 200
DB 1354 AAAAGCCTAAAAAATTAATCTCAATTTATGAGGAGATGCTCTAAGAGAAAAATACATTGGT 1413
QY 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
DB 1414 TCAGAACATGCAACAAGAGACCAAGCTGAAACACAGTGTCAATGAGGAAGCTGACAC 1473
QY 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGlnSerLeuAsp 240
DB 1474 ATGTATCAAAACGACAAAGATAATGTGAACAAACACACTGAACAGCAGAGTCTTAGAT 1533
QY 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
DB 1534 CAGAAATTTATTTCAACTACAAAGCAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1593
QY 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
DB 1594 CATAGAAGAGCTGACACAAAGCAAGATTAACATTTGATTTTCATTTCTTGAGAGGAAA 1653
QY 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluilePheAsnTyrAsnAsnHisLeu 300
DB 1654 ATGCAACATCATCTCTCTAAAGAGAAAAATGAGGAGATATTTAATACAATAACCATTTA 1713
QY 301 LysAsnArgIleTyrGlnTyrGlnLysGluLysAlaGluThrGluAsnSer 317
DB 1714 AAAAACCCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAACTCA 1764

RESULT 10
AAFI7980
ID AAFI7980 standard; cDNA; 2307 BP.
XX
AC AAFI7980;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human breast cancer associated B726P splice sequence #5.
XX
KW Human; breast cancer associated gene; vaccine; diagnosis; therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200060076-A2.
XX
PD 12-OCT-2000.
XX
PF 15-FEB-2000; 2000WO-US005308.
XX
PR 02-APR-1999; 99US-00285480.
XX
PR 23-JUN-1999; 99US-00339338.
XX
PR 03-SEP-1999; 99US-00389681.
XX
PR 03-NOV-1999; 99US-00433826.
XX
PA (CORI-) CORIXA CORP.
XX

```

PI Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
XX WPI; 2001-122627/13.
XX
XX
PT An isolated polypeptide useful for the treatment and diagnosis of tumors
PT e.g. breast cancer comprises at least an immunogenic portion of a breast
PT tumor protein.
XX
XX
PS Claim 6; Page 229-230; 239pp; English.
XX
XX The present invention provides the coding sequences and some protein
CC sequences of proteins associated with breast cancer in humans. These
CC sequences can be used in the diagnosis and treatment of cancers,
CC particularly breast tumors
XX
SQ Sequence 2307 BP; 919 A; 409 C; 439 G; 540 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.13e-136 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-25 (1-317) x AAF17980 (1-2307)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 1066 ATGGGAACAGAGCTCTGACGTGTGAGGTTCTCACATCTCATGAAATGAAATATATCTC 1125

Qy 21 LeuHisGluAsnCysMetLeuLysLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
Db 1126 TTACATGAAATTCATGTTGAAAGGAAATTCGCTTAAACTGGAATGGAATGACCA 1185

Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db 1186 CTGAAACACCAATACCAGGAAAGGAAATTAATACCTTTGAGGACATTAAGATTTTAA 1245

Qy 61 GluLysAsnAlaGluLeuGlnMetThrLysLysLeuLysGluLysGluSerLeuThrLysArg 80
Db 1246 GAAAGAAATGCTGAACTTCAGATGACCTTAAACTGAAAGAGGAATCATTTAACTAAA 1305

Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db 1306 GCATCTCATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGACACAACTCTCACTCT 1365

Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro 120
Db 1366 AAATTGAAGGAAACACAGACAAAGAAATCTAGAGGCAGAAATTTGAATCACACATCCT 1425

Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnIleValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAAGACCATGATCAAAATTTGACATCAAGAAAGTCAAGAA 1485

Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CTGCTTTCCATTCAGAGAGATGCTTTGTCAGAGAAATTAAGATTTGATGTGAGT 1545

Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTAGATATATAACAATGAGGTGCTCCATCAACCCATTTCTGAAGCTCAAGGAAATCC 1605

Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 1606 AAAAGCCCTAAAATTAATCTCAATATATCAGGAGATGCTTAAAGAGAAATATACATGGT 1665

Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1666 TCAGNACATGTCACAAGAGACCAAGTGAACACAGTGTCAATGAAGAGAGCTGAACAC 1725

Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnSerLeuAsp 240

Db 1726 ATGTATCAAAACGAACAAGATAATGTGAACAAACACACTCAACAGCAGGAGTCTCTAGAT 1785
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnLeuValHisAla 260
Db 1786 CAGAAATATTTCACTCAAGCAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1845
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1846 CATAAGAAAGCTGACAAACAAAGCAAGATAACAATTCATATTCATTTCTTGAGAGGAAA 1905
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluLysLeuPheAsnTyrAsnAsnHisLeu 300
Db 1906 ATGCAACATCATCTCTCTAAAGAGAAATGAGAGATATTTAATTAACAATAACATTTA 1965
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAAACCGTATATATCATATATGAAAAGAGAAAGCAGAAACAGAAAACCTCA 2016

RESULT 11
AAI67223
ID AAI67223 standard; cDNA; 2307 BP.
XX
AC AAI67223;
XX
DT 11-FEB-2002 (first entry)
XX
DB B726P splice variant encoding cDNA.
XX
XX Genetic subtraction; DNA microarray analysis; polymerase chain reaction;
KW cancer; B726P; ss.
XX
XX Homo sapiens.
XX
XX WO200175171-A2.
XX
PD 11-OCT-2001.
XX
XX 02-APR-2001; 2001WO-US010631.
XX
PR 03-APR-2000; 2000US-0194241P.
PR 20-JUL-2000; 2000US-0219862P.
PR 27-JUL-2000; 2000US-0221300P.
PR 18-DEC-2000; 2000US-0256592P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Houghton RL, Dillon DC, Molesh DA, Xu J, Zehentner B, Persing DH;
XX WPI; 2001-626449/72.
XX
XX Identifying tissue (tumor)-specific polynucleotides overexpressed in
PT tissue of interest as compared to control tissue, for detecting cancer
PT cells in patient, comprises DNA microarray analysis or quantitative
PT polymerase chain reaction.
XX
PS Claim 4; Page 107-108; 127pp; English.
XX
XX The invention relates to identifying tissue-specific polynucleotides (p)
CC that involves performing a genetic subtraction to identify pool of (p)
CC from tissue of interest (Ti), performing DNA microarray analysis to
CC identify first subset of polynucleotides (SP1) at least 2-fold over
CC expressed in Ti, and performing quantitative polymerase chain reaction
CC (PCR) analysis on SP1 to identify second subset of (p). The method is
CC useful for determining the presence or absence of a cancer cell in a
CC patient, monitoring the progression of cancer in a patient using a
CC biological sample such as blood, serum, lymph nodes, bone marrow, sputum,
CC urine or a tumor biopsy sample. The methods are useful for determining
CC the presence or absence of or monitoring progression of prostate, breast,
CC colon, ovarian, lung, head and neck, lymphoma, leukemia, melanoma, liver,
CC gastric, kidney, bladder, pancreatic or endometrial cancer. Sequences
CC AAI67218-223 represent determined cDNA sequences of splice variants of
XX B726P

SQ Sequence 2307 BP; 919 A; 409 C; 439 G; 540 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3 13e-136 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-25 (1-317) x AAI67223 (1-2307)

QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
DB 1066 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 1125
QY 21 LeuHisGluAsnCysMetLeuLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
DB 1126 TTACATGAAATTTGCATGTTGAAAGGAAATTTGCCATGCTTAAACTGGAAATAGCCACA 1185
QY 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
DB 1186 CTGAACACCAATACAGAAAGGAAATTAATAATCTTTGAGACATTAAGATTTTAA 1245
QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
DB 1246 GAAAGAGTCTGAACTTCAGATGACCTTAAACTGAAAGAGGAATCATTAACCTAAAGG 1305
QY 81 AlaSerGlnTyrSerGlyGlnLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
DB 1306 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGGAACACATGCTCACTTCT 1365
QY 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro 120
DB 1366 AAATTGAAGGAAACACAGAAAGAAATCTAGAGGCAAAATTTGATCACACCATCTCT 1425
QY 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
DB 1426 AGACTGGCTTCTGCTGTACAGACCATGATCAAAATTTGTGACATCAAGAAAAAGTCAAGAA 1485
QY 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
DB 1486 CCTGCTTTCCACATGTGCAGGAGATGCTTTGTCAAAGAAAAATGAATGTTGATGTAGT 1545
QY 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
DB 1546 AGTACATATATAACCATAGGTGCTCCATCAACCACTTTCTGAGCTCAAGAGAAATCC 1605
QY 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
DB 1606 AAAAGCCTAAAAATTAATCTCAATTTATGCGAGAGATGCTTAAGAGAAAAATACATTGGTT 1665
QY 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
DB 1666 TCAGAACATGTCACAAAGAGACCAACGTGAACACACAGTGTCAAATGAAGGAAGCTGAACAC 1725
QY 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGlnSerLeuAsp 240
DB 1726 ATGTATCAAAACGACAAAGATATATGTAACAAACACACTGTAACAGCAGGAGTCTTAGAT 1785
QY 241 GlnLysPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
DB 1786 CAGAAATTTTCACTACAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1845
QY 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
DB 1846 CATAGAAAGCTGACAAACAAAAGCAAGTAACAAATTTGATATTTCTTTCTTGAGAGAAA 1905
QY 281 MetGlnHisIleLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
DB 1906 ATGCAACATCATCTCTCTAAAGAGAAATAGGAGATATTTTATTTACAAATACCAATTTA 1965
QY 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317

DB 1966 AAAAACCGTATATATCAATATGAAAAAGAGAAAGCAGAAACAGAAACTCA 2016

RESULT 12
AAS47410
ID AAS47410 standard; cDNA; 2307 BP.

XX AAS47410;

XX 18-DEC-2001 (first entry)

XX Human cDNA clone 19310.seq_B726P encoding a breast cancer protein.

XX Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy.

OS Homo sapiens.

XX WO200179286-A2.

XX 25-OCT-2001.

XX 12-APR-2001; 2001WO-US012164.

XX 17-APR-2000; 2000US-00551621.

PR 08-JUN-2000; 2000US-00590751.

PR 22-JUN-2000; 2000US-00604287.

PR 20-JUL-2000; 2000US-00620405.

XX (CORI-) CORIXA CORP.

PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
WPI; 2001-611721/70.

PT Breast Tumor Proteins and nucleic acids useful for the prevention,

PS diagnosis and treatment of breast cancer.

XX Claim 1; Page 274-275; 297pp; English.

XX The invention relates to isolated breast tumour proteins and nucleic acids that encode them, including immunogenic fragments of the proteins. Also included are expression vectors expressing the proteins, transformed cells and antibodies raised against the proteins or an antigen presenting cell expressing the protein. The proteins and nucleic acids may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate breast tumour protein expression, i.e. breast tumours and breast cancer e.g. by gene therapy. The nucleic acids and their complements may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The proteins, nucleic acids and antibodies may be used in assays to identify modulators (e.g. antagonists) of breast tumour protein expression and activity. The antibodies and antagonists may also be used to down regulate expression and activity. The antibodies may also be used as diagnostic agents for detecting the presence of the proteins in samples (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immunopurification diagnostic techniques. The present sequence is a cDNA from a breast tumour cDNA library isolated by subtractive hybridisation against a normal breast cDNA library

SQ Sequence 2307 BP; 919 A; 409 C; 439 G; 540 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3 13e-136 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-25 (1-317) x AAS47410 (1-2307)

```
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db |||||
1066 ATGGGACACAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAATGAAATATATCTC 1125

Qy 21 LeuHisGluAsnCysMetLeuLysLysGluLeuAlaMetLeuLysLeuGluIleAlaThr 40
Db |||||
1126 TTACATGAAATTCATGTTGAAAGGAAATGCCATGCTTAAACATGGAATAGCCACA 1185

Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db |||||
1186 CTGAACACCAATACAGGAAAGGAAATAAATACTTTGAGGACATTAAGATTTTAAAA 1245

Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluSerLeuThrLysArg 80
Db |||||
1246 GAAAGAGATGCTGAACCTTCAGATGACCTTAAACTGAAAGAGGAATCATTTAACTTAAAGG 1305

Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db |||||
1306 GCATCTCAATATAGTGGGCAGCTTAAAGTTCGTAGCTGAGAACACAAATGCTCACTTCT 1365

Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPhePro 120
Db |||||
1366 AAATTGAAGGAAACCAAGACAAAGAAATACTAGAGGCAGAAATTTGAATCACACCATCTCT 1425

Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db |||||
1426 AGACTGGCTTCTGCTTACAGACATGATCAATTTGTGACATCAAGNAAAGTCAAGAA 1485

Qy 141 ProLysPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db |||||
1486 CCTGCTTTCCACATTCGACGAGAGTGTGTTGTCAAAGAAAAATGAAATGTTGATGTGAGT 1545

Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db |||||
1546 AGTACGATATATAAACAATGAGGTGCTCCATCAACCACTTTCTGAAGCTCAAGGAAATCC 1605

Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db |||||
1606 AAAAGCCTTAAANNTAATCTCAATATGACGAGATGCTCTTAAGAGAAATACATTTGGTT 1665

Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db |||||
1666 TCAGAACATGTCACAAAGAGACCAACGTCGAAACACACAGTGTCAAAATGAAGGAAGCTGAACAC 1725

Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db |||||
1726 ATGTATCAAAACGAACAAGATAATGTGAAACAACACACACTGAACACAGCAGGAGTCTCTAGAT 1785

Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
Db |||||
1786 CAGAAATTTATTTTCAACTACAAGGCNAAATATGTGGCTTCAACAGCAATTAATGTCATGCA 1845

Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db |||||
1846 CATGAAGAAAGCTGCACAAACAAAGCAAGATAACAATTTGATATTCATTTCTTCTTGAGAGGAAA 1905

Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db |||||
1906 ATGCAACATCATCTCTCTTAAAGAGAGAAATGAGGAGATATTTAATTAATCAATATCAATTTA 1965

Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db |||||
1966 AAAAACCGTATATATCATATATGAAAGAGAAAGAGAGAAAGAGAGAAAGAGAGAAAGAGAAAGTCA 2016

RESULT 13
ID ABS64011
XX ABS64011 standard; cDNA; 2307 BP.
AC ABS64011;
XX
XX 15-NOV-2002 (first entry)
XX
XX Human breast tumour polynucleotide #463.
DE
```

```
XX Human; breast tumour protein; gene; ss; breast cancer; cytostatic;
KW vaccine.
XX Homo sapiens.
XX OS
XX PN US2002085998-A1.
XX
XX PD 04-JUL-2002.
XX
XX PF 13-APR-2001; 2001US-00834759.
XX
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX
XX PA (CORI-) CORIXA CORP.
XX
XX PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI Henderson RA;
XX
XX WIPI; 2002-635657/68.
XX
XX Novel breast cancer polynucleotides and polypeptides encoded by the
XX polynucleotides, useful for detecting the presence of breast cancer in a
XX patient, and in pharmaceutical compositions, for treating breast cancer.
XX
XX Claim 1; Page 200-201; 247pp; English.
XX
XX The invention relates to an isolated breast tumour polynucleotide and the
XX polypeptide it encodes. The polynucleotide and polypeptide are useful for
XX detecting the presence of breast cancer in a patient, and in
XX pharmaceutical compositions for treating breast cancer. The sequences are
XX useful for stimulating an immune response in a patient and can therefore
XX be used in production of vaccines. The sequences are also useful for
XX detecting the presence of a cancer in a patient, by obtaining a
XX biological sample from the patient, contacting the biological sample with
XX a composition of the invention and detecting the amount of polynucleotide
XX that hybridizes to the sample. This sequence represents a human breast
XX tumour polynucleotide of the invention
XX
SQ Sequence 2307 BP; 919 A; 409 C; 439 G; 540 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,13e-136 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x ABS64011 (1-2307)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db |||||
1066 ATGGGACAAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAATGAAATATATCTC 1125

Qy 21 LeuHisGluAsnCysMetLeuLysLysGluLeuAlaMetLeuLysLeuGluIleAlaThr 40
Db |||||
1126 TTACATGAAATTCATGTTGAAAGGAAATTTGCCATGCTTAAACATGGAATAGCCACA 1185

Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db |||||
1186 CTGAACACCAATACAGGAAAGGAAATTAATTAATTTGAGGACATTAAGATTTTAAAA 1245

Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluSerLeuThrLysArg 80
Db |||||
1246 ATGGGACAAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAATGAAATATATCTC 1305
```

```
Db 1246 GAAAGAAATGCTGAACCTTCAGATGACCCCTAAACCTGAAAGAGGAATCAATTAACATAAAGG 1305
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
Db 1306 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGACACACATGCTCACTTCT 1365
Qy 101 LysLeuLysGluLysGlnAspLysGluLeuLeuAlaGluLeuGluSerHisHisPro 120
Db 1366 AAATTGAAGGAAACAAAGACAAAGAAATACTAGAGGAGAGAAATGAATCAACCATCCT 1425
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAAGACCATCATCAATTTGTGACATCAAGAAAGATCAAGAA 1485
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CTGCTTTCCATTTGACAGAGATGCTTTGTCAGAGAAATGAATGTTGATGTGAGT 1545
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTACGATATATAACAATGAGGTGCTCCATCAACCATTTCTGAAGCTCAAGGAAATCC 1605
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 1606 AAAAGCCTAAAAATTAATCTCAATTTATGACGAGATGCTCTAAGAGAGAAATACATTTGTT 1665
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1666 TCAGAACATGCAACAAGAGACCAACGTAACACAGTGTCAATGAGAGAGCTGAACAC 1725
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnSerLeuAsp 240
Db 1726 ATGTATCAAAACGAACAAGATAATGTGAACAAACACACTGAACAGCAGGAGTCTCTAGAT 1785
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
Db 1786 CAGAAATTAATTTCAACTACAAGACAAAATATGTGGCTTCAACAGCAATTAATGTCATGCA 1845
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1846 CATAAGAAGCTGTGACAAACAAAGCAAGATACAAATGATATTCATTTCTTGAGAGGAAA 1905
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluLeuPheAsnTyrAsnAsnHisLeu 300
Db 1906 ATGCAACATCATCTCTCTAAAGAGAGAAATAGGAGATATTTAATTAACAATAACCATTTA 1965
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAAACCGTATATATCAATATGAAAAAGAGAAAGAGAGAAACAGAAACTCA 2016

RESULT 14
ABT33223
ID ABT33223 standard; DNA; 2307 BP.
XX
AC ABT33223;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human tumour-related DNA sequence - SEQ ID No 468.
XX
KW Human; ds; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX
OS Homo sapiens.
XX
PN WO200283956-A1.
XX
PD 24-OCT-2002.
XX
PF 15-APR-2002; 2002WO-US012378.
XX
PR 13-APR-2001; 2001US-00834759.
PR 07-DEC-2001; 2001US-00007805.
```

```
PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
PI Vedvick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
XX New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX Example 1; Page 290-291; 375pp; English.
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumor protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumors (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumor protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present DNA sequence represents a
CC human tumour-related DNA sequence
XX
SQ Sequence 2307 BP; 919 A; 409 C; 439 G; 540 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,13e-136 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-489-079-25 (1-317) x ABT33223 (1-2307)
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 1066 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 1125
Qy 21 LeuHisGluAsnCysMetLeuLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
Db 1126 TTACATGAAATTTGATGATGTTGAAAAGGAAATTTGCGTCTGAAATCTGAAATATGCCACA 1185
Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db 1186 CTGAAACACCAATACCAGGAAAGGAAATTAATAATCTTTGAGGACATTAAGATTTTAAAA 1245
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluSerLeuThrLysArg 80
Db 1246 GAAAAGAAATGCTGAACCTTCAGATGACCTTAAACTGAAAGAGGAATCATTAACATAAAGG 1305
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
Db 1306 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGACACACATGCTCACTTCT 1365
Qy 101 LysLeuLysGluLysGlnAspLysGluLeuLeuAlaGluLeuGluSerHisHisPro 120
Db 1366 AAATTGAAGGAAACAAAGACAAAGAAATACTAGAGGAGAGAAATGAATCAACCATCCT 1425
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAAGACCATCATCAATTTGTGACATCAAGAAAGATCAAGAA 1485
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CTGCTTTCCATTTGACAGAGATGCTTTGTCAGAGAAATGAATGTTGATGTGAGT 1545
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTACGATATATAACAATGAGGTGCTCCATCAACCATTTCTGAAGCTCAAGGAAATCC 1605
```


Db 1846 CATAGAAAGCTGACACAAAGCAAGATAACAATTGATATTCATTTCTTGAGAGGAAA 1905
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnHisLeu 300
Db 1906 ATGCAACATCATCTCTTAAAGAGAAATATGAGGAGATATTTAATTACAATACCATTTA 1965
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAAACCGTATATATCAATATGAAAAAGAGAAAAAGCAGAAACAGAAAACTCA 2016

Search completed: August 1, 2005, 23:18:07
Job time : 603.051 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:05:16 ; Search time 88.6398 Seconds
(without alignments)
1831.334 Million cell updates/sec

Title: US-09-489-079-25

Perfect score: 1632

Sequence: 1 MCTRALQCEVSHTHENENYL.....NHLKNRIYQYEKAETENS 317

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1586	97.2	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	1258	77.1	424	2 Q9NSI9	Q9nsi9 homo sapien
3	819.5	50.2	1080	2 Q6ZRI4	Q6zri4 homo sapien
4	819.5	50.2	1710	2 Q9H1Q1	Q9h1q1 homo sapien
5	819.5	50.2	1715	2 Q9UFS8	Q9ufs8 homo sapien
6	669	41.0	641	2 Q60311	Q60311 homo sapien
7	632	38.7	1043	2 Q69ZS2	Q69za2 mus musculus
8	565	34.6	823	2 Q9H0H6	Q9h0h6 homo sapien
9	556	34.1	292	2 Q68DM0	Q68dm0 homo sapien
10	552.5	33.9	433	2 Q8I2N7	Q8i2n7 homo sapien
11	546.5	33.5	718	2 Q9HCD1	Q9hcd1 homo sapien
12	497	30.5	992	1 AN18 HUMAN	Q8ivf6 homo sapien
13	411	25.2	1011	2 Q9BXX2	Q9bxx2 homo sapien
14	334	20.5	483	2 Q6ZU57	Q6zu57 homo sapien
15	266.5	16.3	119	2 Q96IX9	Q96ix9 homo sapien
16	250.5	15.3	163	2 Q6ZU74	Q6zu74 homo sapien
17	212.5	13.0	291	2 Q8NF67	Q8nf67 homo sapien
18	212	13.0	132	2 Q8TDH5	Q8tdh5 homo sapien
19	202	12.4	187	2 Q8TDH6	Q8tdh6 homo sapien
20	192	11.8	1500	2 Q75240	Q75j40 oryza sativ
21	186.5	11.4	1875	1 MLP1 YEAST	Q02455 saccharomyc
22	183.5	11.2	1294	2 Q6X1Y7	Q6x1y7 legionella
23	183	11.2	887	2 Q6DCA5	Q6dcas homo sapien
24	181.5	11.1	2474	2 Q6RT24	Q6rt24 mus musculus
25	180.5	11.1	1978	1 MYHB CHICK	P10587 gallus gall
26	179.5	11.0	915	2 Q869R0	Q869r0 dictyosteli
27	177.5	10.9	1155	2 Q7XXP7	Q7xxp7 oryza sativ
28	177.5	10.9	1972	1 MYHB RABIT	P35748 oryctolagus
29	176.5	10.8	1972	1 MYHB MOUSE	O08638 mus musculus
30	176.5	10.8	1984	2 Q692X3	Q692x3 mus musculus
31	175	10.7	1738	2 Q76329	Q76329 dictyosteli

32 175 10.7 1781 2 Q86KX8
33 174.5 10.7 1677 2 Q7KT48
34 174.5 10.7 1690 1 C190 DROME
35 174.5 10.7 2090 1 NIN HUMAN
36 173.5 10.6 1005 2 Q81JK0
37 173.5 10.6 1602 2 Q6BFF0
38 173 10.6 555 2 Q9C7V7
39 173 10.6 813 2 Q6FPZ7
40 173 10.6 1374 2 Q86XU5
41 173 10.6 1960 1 MYH9 HUMAN
42 172.5 10.6 724 2 Q7Z659
43 172.5 10.6 1972 2 Q8R384
44 172 10.5 1979 1 TRIA_HUMAN
45 172 10.5 2101 2 Q14981

ALIGNMENTS

RESULT 1

Q9BXX3 PRELIMINARY; PRT; 1341 AA.
AC Q9BXX3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Breast cancer antigen NY-BR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=2117479; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RT Identification of a Tissue-Specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library."
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269087; AAK27325.1; -
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001110; ANK.
DR InterPro; IPR001969; Pept_Asp_As.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKTRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
KW ANK repeat.
SQ SEQUENCE 1341 AA; 152776 MW; 33E53DDE6FD3A58B CRC64;

Query Match 97.2%; Score 1586; DB 2; Length 1341;
Best Local Similarity 99.7%; Pred. No. 1.9e-77;
Matches 308; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 EVSHTHENENYLHENCMLKKEIATLKHQYQEKYKFEDIKILKEKNALQM 68
1033 QVSHHTHENENYLHENCMLKKEIATLKHQYQEKYKFEDIKILKEKNALQM 1092
QY 69 TLKKEESITKRAQSYGOLKVLIAENTMLTSLKKEQDKKEILAEIESHHPRLASAVQD 128
1093 TLKKEESITKRAQSYGOLKVLIAENTMLTSLKKEQDKKEILAEIESHHPRLASAVQD 1152
QY 129 HDQIVTSRKSQEPAPFIAGDACLQRKNVDVSSITYNNEVLHQPPLSPAQRKSKSLKINLN 188
1153 HDQIVTSRKSQEPAPFIAGDACLQRKNVDVSSITYNNEVLHQPPLSPAQRKSKSLKINLN 1212
QY 189 YAGDALRENTLVSEHAQRQORETQCCKEAEHYQNEQDNVNHKGTQOESLDQKLFQLOS 248

```

Db 1213 YAGDALRENTLVSEHAQDQRETQCMKEAEHMYQNEQDNNVKNHTEQOQESLDQKLFQIQS 1272
Qy 249 KNNMLQQQLVHAHKKADNKSKITIDIHFLERKQKHLLKKEKNEEFNNYNNHLKNRIYOYE 308
Db 1273 KNNMLQQQLVHAHKKADNKSKITIDIHFLERKQKHLLKKEKNEEFNNYNNHLKNRIYOYE 1332
Qy 309 KEKASTENS 317
Db 1333 KEKASTENS 1341

RESULT 2
QNSI9
ID Q9NSI9 PRELIMINARY; PRT; 424 AA.
AC Q9NSI9
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE PRE04 protein (Fragment).
GN Name=PRE04;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.S.,
RA Toyoda A., Ishii K., Toki Y., Choi D.K., Soeda E., Ohki M.,
RA Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald A., Rump A., Schillhaber M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis P.,
RA Lebrach H., Reinhardt R., Vaspo M.L.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL163203; CAB90394.1; -.
FT NON TER 1
FT TER 424
SQ SEQUENCE 424 AA; 50038 MW; C6E704795C534625 CRC64;

```

```

Query Match 77.1%; Score 1258; DB 2; Length 424;
Best Local Similarity 82.1%; Pred. No. 2.6e-60;
Matches 252; Conservative 23; Mismatches 30; Indels 2; Gaps 2;

Qy 9 EVSHTHENYLLHNCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKLKEKNAELQM 68
Db 120 QVSHTHSENDLFHNCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKLKEKNAELQM 179
Qy 69 TLKKEESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKLEIAEIESHHPRLASAVOD 128
Db 180 TLKKEESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKLEIAEIESHHPRLASAVOD 239
Qy 129 HDQIVTSRKSQSPAFHAGDAGLQRMKNDVSVSTIYNNEVLHOPLSRQSKSLKINL 188
Db 240 HDQIVTSRKSQSPAFHAGDAGLQRMKNDVSVSTIYNNEVLHOPLSRQSKSKINL 299
Qy 189 YAGDALRENTLVSEHAQDQRETQCMKEAEHMYQNEQDNNVKNHTEQOQESLDQKLFQIQS 248
Db 300 YAGDDLRENALVSEHAQDQRETQCMKEAEHMYQNEQDNNVKNHTEQOQESLDQKLFQIQS 359
Qy 249 KNNMLQQQLVHAHKKADNKSKITIDIHFLERKQKHLLKKEKNEEFNNYNNHLKNRIYOYE 308
Db 360 KNNMLQQQLVHAHKKV-NKSKVTINIQFPETKQKH-LKEKNEEFNNYNNHLKECIDOYE 417
Qy 309 KEKASTE 315
Db 418 KEKASTE 424

```

```

RESULT 3
Q6ZR14
ID Q6ZR14 PRELIMINARY; PRT; 1080 AA.
AC Q6ZR14;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ46736.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK128577; BAC87508.1; -.
FT NON TER 1080
SQ SEQUENCE 1080 AA; 126741 MW; 47BBAAA933AE595C CRC64;

Query Match 50.2%; Score 819.5; DB 2; Length 1080;
Best Local Similarity 54.0%; Pred. No. 3.3e-36;
Matches 177; Conservative 50; Mismatches 88; Indels 13; Gaps 4;

Qy 1 MGTALQCEVSHTEHENYLLHNCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKL 60
Db 448 MAQKNSENSESHSEEEKDLKSHKMLQEEIAMLEIDTIKNQNEKEKCFEDLKIVK 507
Qy 61 EKNAELQWLKLEESLTKRASQYSGQLKVLIAENTMLTSLK- EKQDKLEIAEIESHH 119
Db 508 EKNEEDLQTKIQNEETLTQTISQYNGRLSVLTAEANMLSKLENEKQSKERLEAEVESH 567
Qy 120 PRLASAVODHQIVTSRKSQSPAFHAGDAGLQRMKNDVSVSTIYNNEVLHOPLSRQSKSL 177
Db 568 SRLAAIHDRDQSETSKRELEAFQARDECSRLQDMNFDVSNLKNNEILSQQLFKTE 627
Qy 178 RKSLSKLINLVAGDALRENTLVSEHAQDQRETQCMKEAEHMYQNEQDNNVKNHTEQOQ 237
Db 628 SKLNSLEIEFHTRDALREKTLGLERVQDLISQTCQCKENEQKYQNEQVKNVYIGQE 687
Qy 238 SLQDKLQFQSKNNMLQOQLVHAHKKADNKSKITIDI----HFLERKQKH-----LLK 287
Db 688 SVEERLSQLQSENMLLRQQLDADNADNADNADNADNADNADNADNADNADNADNADNADN 747
Qy 288 EKNEEFNNYNNHLKNRIYOYEKEKASTE 315
Db 748 ERNKELISECNHLKBERQYENKAE 775

RESULT 4
Q9HIQ1
ID Q9HIQ1 PRELIMINARY; PRT; 1710 AA.
AC Q9HIQ1;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE BA145E8.1 (KTA01074).
GN Name=BA145E8.1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bird C.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AL162272; CAC19649.1; -.
 DR HSSP; P20749; 1K1B.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 4.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 4.
 DR PROSITE; PS50088; ANK_REPEAT; 4.
 DR PROSITE; PS50297; ANK_REPEAT; 1.
 KW ANK repeat.
 SQ SEQUENCE 1710 AA; 196409 MW; 01CBP9BADB894872 CRC64;
 Query Match 50.2%; Score 819.5; DB 2; Length 1710;
 Best Local Similarity 54.0%; Pred. No. 5.5e-36;
 Matches 177; Conservative 50; Mismatches 88; Indels 13; Gaps 4;
 QY 1 MGTRALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQKKNKYPEDIKILK 60
 DB 891 MAQKKNSENSHSHEEKDLSHKNMLQEBIAMLRLIEDITIKNQKEKKEKCFEDIKIVK 950
 QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLK-EKQDKEILAEIESHH 119
 DB 951 EKNEELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLK-EKQDKEILAEIESHH 1010
 QY 120 PRLASAVQDHQIVTSRKSSQSPAFHAGDAC--LQRKQNVDSSTTYNNVNLHQLPSEAQ 177
 DB 1011 SRLAAAIHDRDQSETSKRELEAFQARDECRLQDKMNFVSNLKNNEILSQQLFKTE 1070
 QY 178 RKSKSLKINLVAGDALRENTLVSEHAORDQRETQCMKEAHHYQNEQDNNVNHTEQOR 237
 DB 1071 SKLSLEIEFHHTRDALREKTLGLERVQKLSQTCQCMKEQYQNEQVKVNYIKGQE 1130
 QY 238 SLQDKLFQLOSKNNMLQQQLVHAHKADNKSKITIDI----HFLERKMQHH-----LKK 287
 DB 1131 SVVERLSQSENNMLRQQLDDAHNKADNKEKTVINIQQDFAIVQKLQAESEKQSILLE 1190
 QY 288 EKNEEIFNYYNNHLKNRIYQYEKEKAETE 315
 DB 1191 ERNKELISECNHLKERYQYENKAEERE 1218
 RESULT 5
 Q9UP58 PRELIMINARY; PRT; 1715 AA.
 AC Q9UP58;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE KIAA1074 protein (Fragment).
 GN Name=KIAA1074;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9937452; PubMed=10470851;
 RA Kikuno R., Nagase T., Ishikawa K., Hirose M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIV. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."
 RL DNA Res. 6:197-205(1999).
 DR EMBL; AB028997; BAAB3026.2; -.
 DR HSSP; P20749; 1K1B.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 4.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 4.
 DR PROSITE; PS50088; ANK_REPEAT; 4.
 DR PROSITE; PS50297; ANK_REPEAT; 1.
 KW ANK repeat.
 FT NON TER 1 1

SQ SEQUENCE 1715 AA; 196962 MW; A82858F5F5B83203 CRC64;
 Query Match 50.2%; Score 819.5; DB 2; Length 1715;
 Best Local Similarity 54.0%; Pred. No. 5.5e-36;
 Matches 177; Conservative 50; Mismatches 88; Indels 13; Gaps 4;
 QY 1 MGTRALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQKKNKYPEDIKILK 60
 DB 896 MAQKKNSENSHSHEEKDLSHKNMLQEBIAMLRLIEDITIKNQKEKKEKCFEDIKIVK 955
 QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLK-EKQDKEILAEIESHH 119
 DB 956 EKNEELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLK-EKQDKEILAEIESHH 1015
 QY 120 PRLASAVQDHQIVTSRKSSQSPAFHAGDAC--LQRKQNVDSSTTYNNVNLHQLPSEAQ 177
 DB 1016 SRLAAAIHDRDQSETSKRELEAFQARDECRLQDKMNFVSNLKNNEILSQQLFKTE 1075
 QY 178 RKSKSLKINLVAGDALRENTLVSEHAORDQRETQCMKEAHHYQNEQDNNVNHTEQOR 237
 DB 1076 SKLSLEIEFHHTRDALREKTLGLERVQKLSQTCQCMKEQYQNEQVKVNYIKGQE 1135
 QY 238 SLQDKLFQLOSKNNMLQQQLVHAHKADNKSKITIDI----HFLERKMQHH-----LKK 287
 DB 1136 SVVERLSQSENNMLRQQLDDAHNKADNKEKTVINIQQDFAIVQKLQAESEKQSILLE 1195
 QY 288 EKNEEIFNYYNNHLKNRIYQYEKEKAETE 315
 DB 1196 ERNKELISECNHLKERYQYENKAEERE 1223
 RESULT 6
 O60311 PRELIMINARY; PRT; 641 AA.
 AC O60311;
 DT 01-AUG-1998 (T-EMBLrel. 07, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE KIAA0565 protein (Fragment).
 GN Name=KIAA0565;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98290545; PubMed=9628581;
 RA Nagase T., Ishikawa K., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro."
 RL DNA Res. 5:31-39(1998).
 DR EMBL; AB011137; BAA25491.2; -.
 DR NON TER 1 1
 FT NON TER 1 1
 SQ SEQUENCE 641 AA; 75777 MW; F11566F2E6A0B9F4 CRC64;
 Query Match 41.0%; Score 669; DB 2; Length 641;
 Best Local Similarity 46.1%; Pred. No. 2.4e-28;
 Matches 152; Conservative 59; Mismatches 95; Indels 24; Gaps 5;
 QY 1 MGTRALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQKKNKYPEDIKILK 60
 DB 102 MARKKNVSEISHRQKEKDLFHEDCMLQEBIALRLIEDITIKNQKEKKEKCFEDIEAVK 161
 QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLK-EKQDKEILAEIESHH 119
 DB 162 EKNDNLQKIILKEETLTETILQYSGQLNNLTAEKILNLENGKQNKQERLEIEMESYR 221
 QY 120 PRLASAVQDHQIVTSR-----KQEPAFHAGDACLQKKNVDSSTTYNNVNLHQLP 172
 DB 222 CRLAAAVRDCDSQGTARDLKLDFORTQEWVR-----LHDKMKVDSGLQAKNEILSEK 275

```
QY 173 LSEQRKSKLINLVAGDALRENTLVSEHAORDQRETQCOMKEAHEMYQNEQDNVAKH 232
DB 276 LSNAESKINSIQIHUNTRDALGRESILILERVQRLDSQTQCKKETEOMYQIEQSKLKKY 335
QY 233 TQEQSLDQKLQFQSKNWLQQLVHAHKADNKSII--TTDIHF-----LERKQV 282
DB 336 IAKQSEVERLSQLOSENWLLRQQLDDAHKANSQEKTSSTIQDQFHSAAKNLQAESEKQ 395
QY 283 HHLKEKNEEIPNNYNNHLKNRIYQYEKKA 312
DB 396 ILSLQEKNELMDEYNHLKERMDCQEKKA 425

RESULT 7
Q69ZS2
ID Q69ZS2 PRELIMINARY; PRT; 1043 AA.
AC Q69ZS2;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE MKIAA1074 protein (Fragment).
GN Name=mkIAA1074;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaieho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT cDNAs Identified by Screening of Sequences of 500 Mouse KIAA-Homologous
RT Randomly Sampled from Size-Fractionated Libraries."
RL DNA Res. 11:205-218(2004).
DR EMBL; AK171096; BAD32374.1; -.
DR InterPro; IPR010989; C-share.
DR InterPro; IPR009054; Topismrse_insert.
FT NON_TER 1
FT NON_TER 1043
SQ SEQUENCE 1043 AA; 121354 MW; 5F37B1F9AA37B212 CRC64;

Query Match 38.7%; Score 632; DB 2; Length 1043;
Best Local Similarity 43.9%; Pred. No. 4.2e-26;
Matches 144; Conservative 56; Mismatches 92; Indels 36; Gaps 5;

QY 1 MGTALQCEVSHHTEENYLLHNCMLKKEIAMLKLEIATLKHQYQEKYFEDIKILK 60
DB 325 MTQKQMTSEVSVSHEKEDLLHKNQRLQDEAVLRLEMDTIKSHNQEKYKYLEDIKIAN 384
QY 61 ERNAELQMTLKLKEESLTRASQYSGQLVLAENTMLTSK-LKEQDKKEILEAEIESHH 119
DB 385 EKNDNLQRMVKL-----NMLSSKLDNEKQNKERLETQVESPR 421
QY 120 PRIASAVQHDQIVTSRKQSEPAFHITAGDAC--LQKKNVDVSSITYNNEVLHQLPSEAQ 177
DB 422 SRLASALHDHAEITQATRLDEIAFQARDEWFRVKDKNNFDMNSLRNNEVLSSQLSKTE 481
QY 178 RKSKSLKINLVAGDALRENTLVSEHAORDQRETQCOMKEAHEMYQNEQDNVKNKTEQOE 237
DB 482 RKLSLEIEFHHTYKDELREKTLAKHAQRDLQSQTCQCKEVEHMFQDEQKYSKPMGKQE 541
QY 238 SLDDQKLQFQSKNWLQQLVHAHKADNKSITIDI-----HFLERKMQHHL-LK 287
DB 542 STIERLAQLOSENTLLRQLDDAANKAESKDKTIVNIQDFQDVLTRFQAESQRLSLRL 601
QY 288 EKNEEIPNNYNNHLKNRIYQYEKKAETE 315
DB 602 DRNQELVSECShLRELQCYENEKAE 629
```

```
RESULT 8
QSH0H6
ID Q9H0H6 PRELIMINARY; PRT; 823 AA.
AC Q9H0H6;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434A171.
GN Name=DKFZp434A171;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=resticis;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136793; CAB66727.1; -.
DR HSP; Q60778; IOY3.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 823 AA; 93968 MW; 9D2DD5F52C379B2 CRC64;

Query Match 34.6%; Score 565; DB 2; Length 823;
Best Local Similarity 40.7%; Pred. No. 1.3e-22;
Matches 123; Conservative 64; Mismatches 113; Indels 2; Gaps 2;

QY 12 HTHENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKYFEDIKILKEKNAELQMTLK 71
DB 502 HNHEMKGLMDENCILKADIAIRQEICTMKNDNLKENKYLKDIKIVKETAALKEVIK 561
QY 72 LKEESLTRASQYSGQLVLAENTMLTSK-LKEQDKKEILEAEIESHHPLASAVQHD 130
DB 562 LNEEMITETAFRYQOELDLKAENTRLNAELLKEKESKKRLEADIESQSRLLAAAIKSHS 621
QY 131 QIVTSRKQSEPAFHITAGDACLQKKNVDVSSITYNNEVLHQLPSEARQSKSLKINLYA 190
DB 622 ESKVTERNLKALERTDVSVQVEMSSAISKVKAENEFLTEQLSETQIKFNALKDKFCKT 681
QY 191 GDALRENTLVSEHAORDQRETQCOMKEAHEMYQNEQDNVKNKHTEQEESLDQKLQFQSKN 250
DB 682 RDSLRKKSIALETQVNDLSQTQQQTEMKEMYQNAEAKVNNSTGKWNVCVEERICHLOREN 741
QY 251 MWLQOOLVHAHKADNKSIT-IDIHFLERKMQHHLKEKNEEIPNNYNNHLKNRIYQYEK 309
DB 742 AWLVQQLDDVHQKEDHKBIVTNIQRFIESEKKDLVLEKSKKLMNECDHLKESLFQYER 801
QY 310 EK 311
DB 802 EK 803

RESULT 9
Q68DM0
ID Q68DM0 PRELIMINARY; PRT; 292 AA.
AC Q68DM0;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp781D1722.
GN Name=DKFZp781D1722;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```

RN SEQUENCE FROM N.A.
RP TISSUE=Testis;
RG The German cDNA Consortium;
RA Koehler K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749347; CAH18200.1; -
KW Hypothetical protein.
SQ SEQUENCE 292 AA; 34997 MW; FBD10B00B7173F1D CRC64;

Query Match 34.1%; Score 556; DB 2; Length 292;
Best Local Similarity 43.8%; Pred. No. 1.3e-22;
Matches 130; Conservative 59; Mismatches 84; Indels 24; Gaps 4;

QY 26 MLKKEATLKLKLEATLKHQYQKENVFEDIKLKEKNAELQMTLKEESLTKRASQYS 85
DB 1 MLQEEIARLLEIDALKNNWQKEKPFLEDIKAKENDFLQNALMLKEETPATMIFQDS 60
QY 86 GQLKVLIAENTMLTSKLEK-QDKLEABIESHHPLASAVQDHPQIVTSR----- 136
DB 61 IQPEVLKAENEMSSFKLQKDSQNERLERIESYRCRLAAALQDRQSQASKRQVBCDFQ 120
QY 137 KSOEPAFHAGDACLQKKNVDVSSYINNEVLHQPSEAQRKSKIKINLVAGDALRE 196
DB 121 RTREBWFH-----LKEKNFDMNSLKDKNELSEKLSNAENKIRSLKMKWHQTKDALKE 174
QY 197 NTLVSEHAQRDQRETQCKMKEAEHMYQNEQDNVKNKHTEQEESLDQKLFOLOKNNMLQQ 256
DB 175 KTLVSDVDQDLTQSQRRKEIQFQNEEDKSKYIQKQESLUEERLSQIQENLRLRQ 234
QY 257 LVHAHKADNKSITIDI-----HFLERKMQHLLKKEKEEIPFNYNHHLKRI 304
DB 235 LDEAYKKADNQETIINIOPNAIVKQNAQSEESLLEQKNNLINKCDFIKKI 291

RESULT 10
Q81ZM7 PRELIMINARY; PRT; 453 AA.
AC Q81ZM7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Melanoma-associated antigen (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX BRUNO R., d'Orlando O., Altomonte A., Lamaj E., Maio M., Pucillo C.;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF543495; AAN40505.1; -
KW NON TER
SQ SEQUENCE 453 AA; 53140 MW; 81DF80012122BBCA CRC64;

Query Match 33.9%; Score 552.5; DB 2; Length 453;
Best Local Similarity 42.5%; Pred. No. 3.2e-22;
Matches 136; Conservative 57; Mismatches 100; Indels 27; Gaps 8;

QY 9 EVSHTHENYLLHENCMLKEATLKHQYQKENVFEDIKLKEKNAELQ 68
DB 84 QVSETDEKED-LLHENRLMQDEIARLRLKDTIKNQNLK--KYLKDFEIVKVKHEDLQK 140
QY 69 TLKKEESLTKRASQYSGQLKVLIAENTMLTSKL-KKQDKLEABIESHHPLASAVQ 127
DB 141 ALKNGETLAKTIACYSGQLAALTDTENTTLRSKLEKQESRQRLETQMOSYHCHRLNAAR 200
QY 128 DHQIVTSRKSOEPAFHAGDAC--LQKKNVDVSSYINNEVLHQPSEAQRKSKIKI 185
DB 201 DHQDSHKKRDQELAFQGTVDKCRHLQENLNSHL-----ILSLQLSKAESKSRVLKT 253
QY 186 NLNAGDALRENTLVSEHAQRDQRETQCKMKEAEHMYQNEQDNVKNKHTEQEESLDQKLQ 245

RESULT 11
Q9HCD1 PRELIMINARY; PRT; 718 AA.
AC Q9HCD1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE KIAA1641 protein (fragment).
GN Name=KIAA1641;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. the complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046861; BAB13467.1; -
FT NON TER
SQ SEQUENCE 718 AA; 82972 MW; 2AD05ADE70E3C97F CRC64;

Query Match 33.5%; Score 546.5; DB 2; Length 718;
Best Local Similarity 42.2%; Pred. No. 1.1e-21;
Matches 135; Conservative 57; Mismatches 101; Indels 27; Gaps 8;

QY 9 EVSHTHENYLLHENCMLKEATLKHQYQKENVFEDIKLKEKNAELQ 68
DB 349 QVSETDEKED-LLHENRLMQDEIARLRLKDTIKNQNLK--KYLKDFEIVKVKHEDLQK 405
QY 69 TLKKEESLTKRASQYSGQLKVLIAENTMLTSKL-KKQDKLEABIESHHPLASAVQ 127
DB 406 ALKNGETLAKTIACYSGQLAALTDTENTTLRSKLEKQESRQRLETQMOSYHCHRLNAAR 465
QY 128 DHQIVTSRKSOEPAFHAGDAC--LQKKNVDVSSYINNEVLHQPSEAQRKSKIKI 185
DB 466 DHQDSHKKRDQELAFQGTVDKCRHLQENLNSHL-----ILSLQLSKAESKSRVLKT 518
QY 186 NLNAGDALRENTLVSEHAQRDQRETQCKMKEAEHMYQNEQDNVKNKHTEQEESLDQKLQ 245
DB 519 ELHYTGALKKALVFEHVQSELKQSQMOKIERMYKSGYNTMEKICKEQ-----RFCQ 574
QY 246 LQSKNNMLQQLVHAHKADNKSITIDIHF-----LERKMQH--LLKKEKEEIFN 295
DB 575 LKQNNMLQQLDDARNKADNOEKAILNQARCDARVQNLQAECKRHLLEEDNKLVLN 634
QY 296 YNNHLKNRIYQYEKEKAETE 315
DB 635 ELTHSKEKECQYEKEKAERE 654

RESULT 12
AN18 HUMAN STANDARD; PRT; 992 AA.
AC Q81V76; Q72468;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
```

```

DE Ankyrin repeat domain protein 18A.
GN Name=ANKRD18A; Synonyms=KIAA2015;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-349 FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=2338257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB095935; BAC23111.1; ALT INIT.
DR EMBL; BC056266; AAH56266.1; ALT_TERM.
DR Genew; HGNC:23643; ANKRD18A.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS0088; ANK_REPEAT; 4.
KW ANK repeat; Repeat.
FT REPEAT 67 96 ANK 1.
FT REPEAT 100 129 ANK 2.
FT REPEAT 133 162 ANK 3.
FT REPEAT 166 195 ANK 4.
FT REPEAT 199 228 ANK 5.
FT REPEAT 261 288 K -> E (in Ref. 2).
FT CONFLICT 130 130 Missing (in Ref. 2).
FT CONFLICT 288 288 Missing (in Ref. 2).
SQ SEQUENCE 992 AA; 115664 MW; B9B288F087340D9A CRC64;

Query Match 30.5%; Score 497; DB 1; Length 992;
Best Local Similarity 38.8%; Pred. No. 7.6e-19;
Matches 121; Conservative 67; Mismatches 118; Indels 6; Gaps 5;

QY 7 QCEVSHTEHENYLLHNCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAEL 66
DB 310 QPQDSQSGYKKDAMYGNFMLKQDIAMLKEELYATKNDLSLRKEKKYIOEIKSITEINAF 369
QY 67 QMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKL-KEKQDKLEIAEIESHPRLASA 125

370 EKSVRLNEKMITKTIVARYSQQLNDLKAENARLNLSELEKEKINKERLEAEVESLHSLATA 429
126 VQDHDOIVTSRKSQEPAPFIAGDACLOKQNVDSSTIYNNEVLHQLPSEARQKSKSLKI 185
430 INEYNEIV-ERKDLLEVLWRADDVSRHEKMGSNISQLTDKNELJTEQVHKARVKNPTLKG 488
186 NLNVAGDALRENTLVSEHAQDQRETOCOMKEAEHMYONEQDNVNVKHTQEQESLDQKLQF 245
489 KLRETRDALREKTLALGVSQDLDRQAQRIKEMQKMPNGEAKESQSIGKQNSLEIRIQ 548
246 LQSKNMVLQOQLVHAHKADNKSITIDIH--FLERKMOHLLKKEKEEIFYNNHNLKRR 303
549 QELENLLERQLEDARKEGDNK-BIVINIRHDCLENG-KEDLLERNKELMEKYNLYLKEK 606
304 IYQYSEKAEYTE 315
607 LLOCEKEKAERE 618

RESULT 13
Q9BXX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BXX2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061 (2001).
DR EMBL; AF269088; AAK27326.1; -.
DR HSSP; O75832; IUOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS0088; ANK_REPEAT; 1.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON_TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DE438 CRC64;

Query Match 25.2%; Score 411; DB 2; Length 1011;
Best Local Similarity 84.4%; Pred. No. 3.4e-14;
Matches 81; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 9 EVSHTHENYLLHNCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAELQM 68
DB 916 QVSHTESENDFHNCMLKKEIAMLKLEIATLKHOYQEKENKYPEDIKILKEKNAELQM 975
QY 69 TLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKE 104
DB 976 TLKLKQKTVTKRASQYRQQLKVLTAENTMLTSKLKE 1011

RESULT 14
Q6ZU57 PRELIMINARY; PRT; 483 AA.
AC Q6ZU57;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ43983.
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN NCBI_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Ohnina A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isegai T.,
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK125971; BAC86369.1; -
SQ SEQUENCE 483 AA; 56377 MW; EE47AEB54E1327E1 CRC64;

Query Match 20.5%; Score 334; DB 2; Length 483;
Best Local Similarity 45.8%; Pred. No. 2.1e-10;
Matches 77; Conservative 27; Mismatches 54; Indels 10; Gaps 2;

QY 155 MNVDVSTIYNNVHPLSEAOAKSKSLKINLNYAGDALRENTLVSEHAQRDQRETQCO 214
Db 3 MKVDMGLQAKNHLSEKLSNAESKINSLOIQLEHTFDALGRESLILERVQRLSQTQCO 62

QY 215 MKEAEHMYQNEQDNVKNHTEQESLDQKLPQLOSKNMMLQOQLVHAHKADNKSJI--TI 272
Db 63 KXETEYQYEQSKLKKYIAQSSVEERLSQLOSENMLRLQQLDDAHKANSQEKTSSTI 122

QY 273 DIHF-----LERKQHHLLKKEKEIEFNYNHLLKRIYQYEKEKA 312
Db 123 QDOFHSAAKNLRAESEKQILSQLOFKNKELMDYNNHLKERMDQCEKEKA 170

RESULT 15
Q96IX9 PRELIMINARY; PRT; 119 AA.

AC Q96IX9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE MGC12538 protein.
GN Name=MGC12538;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toohiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007072; AAH07072.1; -
SQ SEQUENCE 119 AA; 14172 MW; E06449B08455E397 CRC64;

Query Match 16.3%; Score 266.5; DB 2; Length 119;
Best Local Similarity 53.6%; Pred. No. 2e-07;
Matches 60; Conservative 17; Mismatches 32; Indels 3; Gaps 2;

QY 1 MGTALQCEVSVTHNENYLLHNCMLKKEIAMLKLEIATLKHOYOEKENKYFEDIKILK 60
Db 1 MGTRTLQPEISDSHEKEEDLLHKNHLMQDEIARLRLEIHTIKQILEK--KYLKDIKILK 58

QY 61 EKNAELQMTLKLKEESLTKRASQYSQQLKVLIAENTMLTSLK-KEKQDKEL 111
Db 59 RKHEDLQKALKQNGEKSTKTIAHYSQLTALTDTENTMLRSKLEKEKQSRQL 110

Search completed: August 1, 2005, 22:22:14
Job time : 91.6398 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:08:06 ; Search time 19.5308 Seconds
(without alignments)
1561.672 Million cell updates/sec

Title: US-09-489-079-25
 Perfect score: 1632
 Sequence: 1 MGTRALQCEVSHTHENYL.....NHLKTRIYQYEKKAETENS 317

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :      PIR_79:*
1:  pir1:*
2:  pir2:*
3:  pir3:*
4:  pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	186.5	11.4	1875	2	S38173	myosin-like protei
2	180.5	11.1	1979	1	S03166	myosin heavy chain
3	177.5	10.9	1972	1	A41604	myosin heavy chain
4	176.5	10.8	1938	2	JCS421	smooth muscle myos
5	176.5	10.8	1972	2	JCS420	smooth muscle myos
6	175	10.7	1738	2	T14867	interaptin - slime
7	174.5	10.7	1690	2	T13030	microtubule bindin
8	173	10.6	555	2	C96667	unknown protein, 7
9	172	10.5	761	2	T03719	probable thyroid r
10	172	10.5	2101	2	A42184	nuclear mitotic ap
11	170	10.4	1959	1	A33977	myosin heavy chain
12	170	10.4	1961	1	A61231	myosin heavy chain
13	168.5	10.3	1939	2	T18372	repeat organellar
14	167	10.2	1999	1	S23801	myosin heavy chain
15	166	10.2	750	2	T38435	coiled coil protei
16	166	10.2	1130	2	T34081	hypothetical prote
17	165.5	10.1	1992	2	A47297	myosin heavy chain
18	165.5	10.1	2007	1	B43402	myosin heavy chain
19	164	10.0	1979	2	C71622	hypothetical prote
20	163.5	10.0	980	2	E71606	hypothetical prote
21	163	10.0	631	2	JC4298	hyaluronan recepto
22	161.5	9.9	944	2	S26710	spindle pole body
23	161.5	9.9	1300	2	I53799	CGI protein - huma
24	161.5	9.9	1356	2	S32763	kinectin 1 - huma
25	161.5	9.9	1964	2	A53282	nonmuscle myosin I
26	161.5	9.9	3225	2	I52300	giantin - human
27	161.5	9.9	3259	1	A56539	giantin - human
28	160	9.8	2663	1	S28261	centromere protein
29	159.5	9.8	1939	1	A46762	myosin alpha heavy

30	159	9.7	1410	1	A57013	early endosome ant
31	159	9.7	2094	2	S33124	tptr protein - huma
32	158.5	9.7	1976	2	A59252	myosin heavy chain
33	158.5	9.7	2168	2	T30171	ninein - mouse
34	158.5	9.7	2331	2	T25410	hypothetical prote
35	155.5	9.5	1033	2	S73693	MG328 homolog P01_
36	155.5	9.5	1388	2	S70633	serine/threonine-s
37	154.5	9.5	2954	2	T14156	kinesin-related pr
38	154	9.4	1790	2	S67593	transport protein
39	153.5	9.4	497	2	H70168	hypothetical prote
40	153.5	9.4	756	2	C64236	protein V (fcrv) h
41	153	9.4	746	2	T47237	myosin II heavy ch
42	153	9.4	1188	2	E89896	chromosome segrega
43	153	9.4	1957	2	T38077	hypothetical coile
44	152.5	9.3	1727	2	T50073	myosin-like coiled
45	152	9.3	1091	2	T34107	hypothetical prote

ALIGNMENTS

RESULT 1

S38173
myosin-like protein MLP1 - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: protein YKR095w; protein YKR415
C;Species: *Saccharomyces cerevisiae*
C;Date: 03-May-1994 #sequence revision 03-May-1994 #text_change 09-Jul-2004
S;Accession: S38173; SA0647; S31207
R;Baladron, V.; Ballesta, J.P.G.; Bou, G.; del Rey, F.; Esteban, P.F.; Garcia-Cantalejo,
submitted to the Protein Sequence Database, March 1994
A;Reference number: S38158
A;Accession: S38173
A;Molecule type: DNA
A;Residues: 1-1875 <BL>
A;Cross-references: UNIPROT:Q02455; EMBL:Z28320; NID:g486586; PID:g486587; MIPS.YKR095W
A;Experimental source: strain S288C
R;Bou, G.; Esteban, P.F.; Baladron, V.; Gonzalez, G.A.; Cantalejo, J.G.; Remacha, M.; Ji
Yeast 9, 1349-1354, 1993
A;Title: The complete sequence of a 15 820 bp segment of *Saccharomyces cerevisiae* chromo
A;Reference number: SA0644; MUID:94205265; PMID:8154186

A;Molecule type: DNA
A;Residues: 1-1875 <BOU>
A;Cross-references: EMBL:X73541; NID:G450550; PIDN:CAA51948.1; PID:G450554
A;Experimental source: strain S288C
R;Koelling, R.; Nguyen, T.; Chen, E.Y.; Botstein, D.
Mol. Gen. Genet. 237, 359-369, 1993
A;Title: A new yeast gene with a myosin-like heptad repeat structure.
A;Reference number: S31207; MUID:93247549; PMID:8483450
A;Accession: S31207
A;Molecule type: DNA
A;Residues: 1-300, 'A', 302-1875 <KOE>
A;Cross-references: EMBL:L01992; NID:G171958; PIDN:AAA34783.1; PID:G171959
C;Genetics:
A;Gene: SGD:MLP1
A;Cross-references: SGD:S0001803; MIPS:YKR095W
A;Map position: 11R

	Query Match	11.4%	Score 186.5	DB 2	Length 1875
	Best Local Similarity	25.1%	Pred. No. 0.0077		
	Matches	82	Conservative	63	Mismatches 127; Indels 55; Gaps 13
Qy	6	LCQEVSHTHENVYLHENCMLKKEATMLKLEIATLKHQYQEKYKFFEDIKILKEKNAE	65		
Db	1226	LRNELENNKKKELOSELDKLKNVAPIESLTAIKYSQEK-----QBLKAKE---E	1278		
Qy	66	LQMTLKLKBSLTKRASQVSGQLKVIATNMTLSKLKEK-QDKEILEAIEHSHPRLAS	124		
Db	1279	VHRWKKRSQDILEKHEQLSSDYKELESE-----LENLKELENKERQGAEEKFNLR	1334		
Qy	125	AVQDHQIVTSRKSQEPAFHAGDACLQRKQNVVDVSTIYNNVLHQPILSEAQRKSKSLK	184		
Db	1335	QAO--PRTKTSKLSD-----SLTEGVNSLRDAKNVLNSLESNARIEELQ	1379		

QY 185 INLVAG-----DALRENTLYSEHAQRDQRETQCOMKEAHEMYQNEQDNVN-----KH 232
Db 1380 -NAKVAQNNQLEAIRK---LQEDAKASRELQAKLESTTSYSTINGLNEETITLKEE 1435
QY 233 TEQESLQDKLFQLOSKNMWQQQLVHAHKKADNKSKITIDIHFLERKMQHHLKKEE 292
Db 1436 LEKQRIQQLOQTSANSQNDLSNIVESMKKSFEEDK---IKFKIKETQ-----EYNEK 1486
QY 293 IFNVNHL---KNRIYQYEKEAETEN 316
Db 1487 ILEAQERLNQPSNINMEIKKKWESEH 1513
RESULT 2
S03166
myosin heavy chain, gizzard smooth muscle [similarity] - chicken
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Gallus gallus (chicken)
C>Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S03166; A27066; A28045; A36604; A43298
R:Yanagisawa, M.; Hamada, Y.; Katsuragawa, Y.; Imamura, M.; Mikawa, T.; Masaki, T.
J. Mol. Biol. 198, 143-157, 1987
A:Title: Complete primary structure of vertebrate smooth muscle myosin heavy chain deduced from complementary DNA
A:Reference number: S03166; MUID:88118918; PMID:2892941
A:Accession: S03166
A:Molecule type: mRNA
A:Residues: 1-1979 <YAN>
A:CROSS-references: UNIPROT:P10587; EMBL:X06546; NID:G63633; PIDN:CAA29793.1; PID:G63634
A:Note: part of this sequence was confirmed by protein sequencing
R:Maiba, T.; Onishi, H.; Yajima, E.; Matsuda, G.
J. Biochem. 102, 133-145, 1987
A:Title: Amino acid sequence of the amino-terminal 24 kDa fragment of the heavy chain of chicken gizzard smooth muscle myosin
A:Reference number: A27066; MUID:88032919; PMID:3312184
A:Accession: A27066
A:Molecule type: protein
A:Residues: 2, 'Z', 4-204 <MAI>
R:Onishi, H.; Maiba, T.; Mivanishi, T.; Watanabe, S.; Matsuda, G.
J. Biochem. 100, 1433-1447, 1986
A:Title: Amino acid sequence of the 203-residue fragment of the heavy chain of chicken gizzard smooth muscle myosin
A:Reference number: A26045; MUID:87194651; PMID:3571180
A:Accession: A26045
A:Molecule type: protein
A:Residues: 653-855 <ONI>
R:Onishi, H.; Maiba, T.; Matsuda, G.; Fujiwara, K.
J. Biol. Chem. 265, 19362-19368, 1990
A:Title: Lys-65 and Glu-168 are the residues for carbodiimide-catalyzed cross-linking between heavy chain monomers of chicken gizzard smooth muscle myosin
A:Reference number: A36604; MUID:91035476; PMID:1977747
A:Accession: A36604
A>Status: preliminary
A:Molecule type: protein
A:Residues: 54-67;146-183 <ON2>
R:Cole, D.G.; Yount, R.G.
Biochemistry 31, 6186-6192, 1992
A:Title: Stability and photochemical properties of vanadate-trapped nucleotide complexes of chicken gizzard smooth muscle myosin
A:Reference number: A43298; MUID:92329440; PMID:1385724
A:Accession: A43298
A>Status: preliminary
A:Molecule type: protein
A:Residues: 169-183 <COL>
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; blocked amino end; coiled coil; hydrolase; methylated amino end
F:87-777/Domain: myosin motor domain homology <WMOT>
F:177-184/Region: nucleotide-binding motif A (P-loop)
F:565-578/Region: actin binding #status predicted
F:639-653/Region: actin binding #status predicted
F:850-1940/Domain: coiled coil <COI>
F:850-1290/Region: S2
F:1291-1979/Region: light meromyosin
F:1941-1979/Domain: carboxyl-terminal <CBT>
F:12/Modified site: blocked amino end (Ser) (in mature form) #status experimental
F:128/Modified site: N6,N6,N6-trimethyllysine (Lys) #status experimental
F:183/Binding site: ATP (Lys) #status predicted

F:707,717/Active site: Cys #status predicted

Query Match 11.1%; Score 180.5; DB 1; Length 1979;
Best Local Similarity 23.4%; Pred. No. 0.018;
Matches 86; Conservative 62; Mismatches 145; Indels 75; Gaps 13; .

QY 15 ENENYLLHENCMLKKEIAMLKLEIATLKHQVQKE-----NKYFEDIK 57

Db 1067 EGESDLHE-----QIAELQAQIAELKAQLAKKEBELQAALARLEDETSKNNALKKIR 1120

QY 58 ILKKNALQMTLKLKBESLTKRASQ---YSGQLKVLIAE-----NTMLTSKLKEKODK 108

Db 1121 ELESISDLQEDLESEKAARNKAERKRDISEEELAKTELEDLTDTATQOELRAKEQ 1180

QY 109 EI-----LEAIESHPRLASAVODHDIIVTSRKSPAPHIAGDAQLQKKNVDVSST 162

Db 1181 ETVVLKRALBEETRTHEAQVQEMQKHTQAVEELTEQLQPKRAKANLDKTKQLEKDNA 1240

QY 163 IYNNEV--LHQPILSEAQRKSKSLKINL----NYAGDALRENTLVSEHAQRDQRETQ---C 213

Db 1241 DLANEISLSQAQKODVEHKKKLEVLQDLQSKYSDGERVTELNEKVHKLQIEVENVTS 1300

QY 214 QMKAE-----HMYQNEQNVNKHTEQESLQDKLFQLOSKNMWQQQL--- 257

Db 1301 LLNEAESKNIKLTQDVATLGSQLODTQELLQETROKLNVTTKLQLEDDEKNSLQEQULDE 1360

QY 258 -VHAHKADNK-SKITIDIHFLERKMQHLL-----KEKNEEIFNVNHLKNRIYQ 306

Db 1361 EVEAKQNLERHISTLTQLSDSKKKLQEFATVETMBEGKKLQOREIESLTQQFEERKAAS 1420

QY 307 YEK-EKAE 313

Db 1421 YDKLEKTK 1428

RESULT 3

A41604

myosin heavy chain, smooth muscle, long splice form - rabbit

N:Contains: myosin ATPase (EC 3.6.4.1)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002

C:Accession: A41604; A33501

R:Babij, P.; Kelly, C.; Periasamy, M.

Proc. Natl. Acad. Sci. U.S.A. 88, 10676-10680, 1991

A:Title: Characterization of a mammalian smooth muscle myosin heavy-chain gene: complete

A:Reference number: A41604; MUID:92073350; PMID:1961735

A:Accession: A41604

A:Molecule type: mRNA

A:Residues: 1-1972 <BAB>

A:CROSS-references: GB:M77812

R:Nagai, R.; Kuro-o, M.; Babij, P.; Periasamy, M.

J. Biol. Chem. 264, 9734-9737, 1989

A:Title: Identification of two types of smooth muscle myosin heavy chain isoforms by cDN

A:Reference number: A33501; MUID:89255535; PMID:2722872

A:Accession: A33501

A:Molecule type: mRNA

A:Residues: 1455-1972 <NAG>

A:CROSS-references: GB:J04833; NID:G165519; PIDN:AAA31407.1; PID:G165520

A:Experimental source: smooth muscle

A:Note: examination by Southern blotting for the regions of difference between this isofo

ative splicing

C:Superfamily: myosin heavy chain; myosin motor domain homology

C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated

F:88-771/Domain: myosin motor domain homology <WMOT>

F:178-185/Region: nucleotide-binding motif A (P-loop)

F:559-572/Region: actin binding #status predicted

F:633-647/Region: actin binding #status predicted

F:844-1938/Domain: coiled coil #status predicted <COI>

F:844-1284/Region: S2

F:1285-1972/Region: light meromyosin

F:1939-1972/Domain: carboxyl-terminal <CBT>

F:123/Modified site: N6,N6,N6-trimethyllysine (Lys) #status predicted

F:184/Binding site: ATP (Lys) #status predicted

F:701,711/Active site: Cys #status predicted

```
Query Match          10.9%; Score 177.5; DB 1; Length 1972;
Best Local Similarity 21.5%; Pred. No. 0.026;
Matches 81; Conservative 66; Mismatches 141; Indels 89; Gaps 12;

QY 4 RALQCEVSHTHENENYLLHNCMLKKEIATLKLHQAQVQDHDQIYTSRKSQBPAPHIAGDACL 63
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1058 RMDCEASDLHE-----QIADLQAQIAELKQWLAKKEEELQAALARLDEI 1103
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 AELQMTLK-----LKESLTKRASQVSG-----QLKVLIAE-----NTM 97
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1104 AQNNALKKIRELEGHISDLQDLDSERAARNKAQKQKRDGLGELEALKTELEDTLDTA 1163
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 98 LPSKLEKQDKKEI-----LEAIESHHPRLASAVQDHDQIYTSRKSQBPAPHIAGDACL 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1164 TQQLRAKREQEVTVLKKALDETRSHQAQVQEMRQKHTQVVEELTEQLQFKRA----- 1218
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 QRKMNVDSSTIYNNE-----VLHQP LSEAQRKSKSLKINL-----NYAGDALRENT 198
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1219 --KANLDKTKQTLKENADLAGELRVLGQAQVEHKKKKLQVQLQELQSKCSDGERARA 1276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 LVSEHAQRDQRTQC---QMKAE-----HMYQNEQDNVKNKHTQEQLSDQ 241
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1277 ELNDKVKHLQNEVESVTGMLSEAGKAIKLAKEVASLGSQLODQTBELLQEETROKLNST 1336
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KLFLQSKNMWLQOOLVHAHKKADN---KSKITIDIHFLERKMQH---HLLKKNKEE 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1337 KLRQLEDBRNSLDQDLDEMEAKQNLERHVSITNLQSDSKKKLQDPFASTIEVMEEGKKR 1396
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 IFNYNHLKNRIYQYEKKA 312
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1397 L---QKMEGLSQYEEKAA 1413
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
JC5421
smooth muscle myosin heavy chain 2 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: JC5421
R:Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A:Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A:Reference number: JC5420, MUID:97242182; PMID:9125171
A:Accession: JC5421
A:Molecule type: mRNA
A:Residues: 1-1938 <HAS>
A:Cross-references: UNIPROT:O08638; DDBJ:D85924; NID:g1945079; PIDN:BAA19691.1; PID:g194
A:Experimental source: smooth muscle
C:Comment: This protein plays a role in smooth muscle cell contraction.
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F:88-771/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)

Query Match          10.8%; Score 176.5; DB 2; Length 1938;
Best Local Similarity 22.4%; Pred. No. 0.029;
Matches 85; Conservative 62; Mismatches 138; Indels 95; Gaps 13;

QY 4 RALQCEVSHTHENENYLLHNCMLKKEIATLKLHQAQVQDHDQIYTSRKSQBPAPHIAGDACL 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1058 RMDCEASDLHE-----QIADLQAQIAELKQWLAKKEEELQAALARLDEI 1103
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 AELQMTLK-----LKESLTKRASQVSG-----QLKVLIAE-----NTM 97
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1104 AQNNALKKIRELEGHISDLQDLDSERAARNKAQKQKRDGLGELEALKTELEDTLDTA 1163
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 98 LPSKLEKQDKKEI-----LEAIESHHPRLASAVQDHDQIYTSRKSQBPAPHIAGDACL 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1164 TQQLRAKREQEVTVLKKALDETRSHQAQVQEMRQKHTQVVEELTEQLQFKRA----- 1218
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 QRKMNVDSSTIYNNE-----VLHQP LSEAQRKSKSLKINL-----NYAGDALRENT 198
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1219 --KANLDKTKQTLKENADLAGELRVLGQAQVEHKKKKLQVQLQELQSKCSDGERARA 1276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 LVSEHAQRDQRTQC---QMKAE-----HMYQNEQDNVKNKHTQEQLSDQ 241
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1277 ELNDKVKHLQNEVESVTGMLSEAGKAIKLAKEVASLGSQLODQTBELLQEETROKLNST 1336
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KLFLQSKNMWLQOOLVHAHKKADN---KSKITIDIHFLERKMQH---HLLKKNKEE 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1337 KLRQLEDBRNSLDQDLDEMEAKQNLERHVSITNLQSDSKKKLQDPFASTIEVMEEGKKR 1396
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 IFNYNHLKNRIYQYEKKA 312
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1397 L---QKMEGLSQYEEKAA 1413
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
T14867
interaptin - slime mold (Dictyostelium discoideum)
```

```
Db 1219 --KANLDKTKQTLKENADLAGELRVLGQAQVEHKKKKLQVQLQELQSKCSDGERARA 1276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 LVSEHAQRDQRTQC---QMKAE-----HMYQNEQDNVKNKHTQEQLSDQ 241
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1277 ELNDKVKHLQNEVESVTGMLSEAGKAIKLAKEVASLGSQLODQTBELLQEETROKLNST 1336
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KLFLQSKNMWLQOOLVHAHKKADN---KSKITIDIHFLERKMQH---HLLKKNKEE 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1337 KLRQLEDBRNSLDQDLDEMEAKQNLERHVSITNLQSDSKKKLQDPFASTIEVMEEGKKR 1396
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 IFNYNHLKNRIYQYEKKA 312
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1397 L---QKMEGLSQYEEKAA 1413
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 5
JC5420
smooth muscle myosin heavy chain 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 09-Jul-2004
C:Accession: JC5420
R:Hasegawa, K.; Arakawa, E.; Oda, S.; Matsuda, Y.
Biochem. Biophys. Res. Commun. 232, 313-316, 1997
A:Title: Molecular cloning and expression of murine smooth muscle myosin heavy chains.
A:Reference number: JC5420, MUID:97242182; PMID:9125171
A:Accession: JC5420
A:Molecule type: mRNA
A:Residues: 1-1972 <HAS>
A:Cross-references: UNIPROT:O08638; DDBJ:D85923; NID:g1945077; PIDN:BAA19690.1; PID:g194
A:Experimental source: smooth muscle
C:Comment: This protein plays a role in smooth muscle cell contraction.
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: nucleotide binding; P-loop
F:88-771/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)

Query Match          10.8%; Score 176.5; DB 2; Length 1972;
Best Local Similarity 22.4%; Pred. No. 0.029;
Matches 85; Conservative 62; Mismatches 138; Indels 95; Gaps 13;

QY 4 RALQCEVSHTHENENYLLHNCMLKKEIATLKLHQAQVQDHDQIYTSRKSQBPAPHIAGDACL 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1058 RMDCEASDLHE-----QIADLQAQIAELKQWLAKKEEELQAALARLDEI 1103
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 AELQMTLK-----LKESLTKRASQVSG-----QLKVLIAE-----NTM 97
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1104 AQNNALKKIRELEGHISDLQDLDSERAARNKAQKQKRDGLGELEALKTELEDTLDTA 1163
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 98 LPSKLEKQDKKEI-----LEAIESHHPRLASAVQDHDQIYTSRKSQBPAPHIAGDACL 151
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1164 TQQLRAKREQEVTVLKKALDETRSHQAQVQEMRQKHTQVVEELTEQLQFKRA----- 1218
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 152 QRKMNVDSSTIYNNE-----VLHQP LSEAQRKSKSLKINL-----NYAGDALRENT 198
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1219 --KANLDKTKQTLKENADLAGELRVLGQAQVEHKKKKLQVQLQELQSKCSDGERARA 1276
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 199 LVSEHAQRDQRTQC---QMKAE-----HMYQNEQDNVKNKHTQEQLSDQ 241
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1277 ELNDKVKHLQNEVESVTGMLSEAGKAIKLAKEVASLGSQLODQTBELLQEETROKLNST 1336
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KLFLQSKNMWLQOOLVHAHKKADN---KSKITIDIHFLERKMQH---HLLKKNKEE 292
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1337 KLRQLEDBRNSLDQDLDEMEAKQNLERHVSITNLQSDSKKKLQDPFASTIEVMEEGKKR 1396
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 IFNYNHLKNRIYQYEKKA 312
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1397 L---QKMEGLSQYEEKAA 1413
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
T14867
interaptin - slime mold (Dictyostelium discoideum)
```


Db 338 VDLEETVSLNEVERKGDIESLMKMSNIEVKRLSNQKLRVTVEQVLTKEGELKRIE 397
Qy 268 SKITIDHFLERKM-----QHLLKKEKNEEFNHNHKNRIYQYEKEKARTEN 316
Db 398 AKHLEEQALLEKIAITHETVYRGLIKEISERV---DSTILNR-FQSLSEKLEEKH 448

RESULT 9
T03719
probable thyroid receptor interactor - human (fragment)
A:Alternate names: CEV14 protein
C:Species: Homo sapiens (man)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C:Accession: T03719
R:Abe, A.; Emi, N.; Tanimoto, M.; Terasaki, H.; Marunouchi, T.; Saito, H.
Blood 90, 4271-4277, 1997
A:Title: Fusion of the platelet-derived growth factor receptor beta to a novel gene CEV14
A:Reference number: Z15027; MUID:98043615; PMID:9373237
A:Accession: T03719
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-761 <AB>
A:Cross-references: UNIPROT:Q15643; EMBL:AF011368; NID:g2618824; PIDN:AA84386.1; PID:g2618824

Query Match 10.5%; Score 172; DB 2; Length 761;
Best Local Similarity 22.6%; Pred. No. 0.017;
Matches 91; Conservative 64; Mismatches 111; Indels 136; Gaps 17;

Qy 16 NENYLLHNCM-LKKEIAMLKLEIATLKHQ-----YQEKENKYFEDIKILKKNVLEIQ 67
Db 248 NENLLRQAVTNLKERILLLEMDIGKJGENEKIVETVYRGKETEY---QALQETNMKFS 303
Qy 68 MTLKLK-----ESLTKRASQ-VSQGLKVLII----- 92
Db 304 MMLREKFECHSMKSKALAFQQLLKEKGQGTGELNQLLNNAVKSMQKTVVFOQERDQVM 363
Qy 93 -----AENTMLTSKLKEKQDEI-----LE-----AETSHHPRLASAVQD----- 128
Db 364 LALKQKQMENTALQNEVQLRDKFESNQELERLNHLLSEDSVTREALAAEDREAKLR 423
Qy 129 -----HDOIVTSRKSQBPAPHIAGDACLQRQNVVDVSTIYNNEVLHQPISLSEAKRS 182
Db 424 KKVTVLEEKLVSSSNAMENASHQASVQVESLQELQNVSVKQRDETALQLSVSQEQVKQYA 483
Qy 183 LKI-NLNYAGDALRENTLVSEHAQDOR-----ETOCOM-----KRAEHM----- 221
Db 484 LSLANLQ-----MVLEHFQEEKAMYSAELEKQKQLTAEWKQKQARENLEGKVISLQ 533
Qy 222 -----YQNEQDNVNHKTEQESLDQKLFOQSKNMWLOQQLVHAHK-----ADNKSKITID 273
Db 534 ECLDEANAALDSASRLTEQDLVKEBQIBELKQNELRQEMLDQVQKMLSLANSSEKVD 593
Qy 274 I-----HFLERKMQHH-----LLKEKNEEIPN 295
Db 594 KVLNRNLFGHFTPKQNRHEVRLMLGMSILGVRREEMQLPH 635

RESULT 10
A42184
nuclear mitotic apparatus protein NuMA - human
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A42184; S23376; S55331; S23647; S24554
R:Compton, D.A.; Szilak, I.; Cleveland, D.W.
J. Cell Biol. 116, 1395-1408, 1992
A:Title: Primary structure of NuMA, an intranuclear protein that defines a novel pathway
A:Reference number: A42184; MUID:92176238; PMID:1541636
A:Accession: A42184
A:Molecule type: mRNA
A:Residues: 1-2101 <COM>
A:Cross-references: UNIPROT:Q14981; EMBL:Z11584; NID:g35120; PIDN:CAAY77670.1; PID:g35121
A:Note: sequence extracted from NCBI backbone (NCBIN:85755, NCBI:85760)

R:Tang, T.K.; Tang, C.C.; Chen, Y.L.; Wu, C.W.
J. Cell Sci. 104, 249-260, 1993
A:Title: Nuclear proteins of the bovine esophageal epithelium. II. The NuMA gene gives rise to two distinct protein products.
A:Reference number: S33376; MUID:93280231; PMID:8505359
A:Accession: S33376
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1705-2101 <TAN>
A:Cross-references: EMBL:Z14229; NID:g296118
A:Note: This translation is not annotated in GenBank entry HSNUMAT3G.. release 113.0
R:Harborth, J.; Weber, K.; Osborn, M.
EMBO J. 14, 2447-2460, 1995
A:Title: Epitope mapping and direct visualization of the parallel, in-register arrangement of the mammalian NuMA.
A:Reference number: S55331; MUID:95300777; PMID:7781599
A:Accession: S55331
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 247-279 <HAR>
R:Yang, C.H.; Lambie, E.J.; Snyder, M.
J. Cell Biol. 116, 1303-1317, 1992
A:Title: NuMA: an unusually long coiled-coil related protein in the mammalian nucleus.
A:Reference number: S23647; MUID:92176231; PMID:1541630
A:Accession: S23647
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-771; Q, 773-814, 'ER', 817-872, 'E', 874-1267, 'R', LRLQAEATSNASARAERSALREEVQSRLR
A:Cross-references: EMBL:Z11583; NID:g35118; PIDN:CAAY77669.1; PID:g35119
A:Note: the authors translated the codon GAG for residue 781 as Gly, TTC for residue 1761
C:Genetics:
A:Gene: GDB:NUMA1; NuMA
A:Cross-references: GDB:137229; OMIM:164009
A:Map position: 11q13-11q13
C:Keywords: mitosis; nucleus

Query Match 10.5%; Score 172; DB 2; Length 2101;
Best Local Similarity 24.2%; Pred. No. 0.057;
Matches 80; Conservative 58; Mismatches 130; Indels 62; Gaps 11;

Qy 6 LQCEVSH-----THENENYLLHNCMLKKEIAMLKLEIATLKHQYQEKENKYPEDIKILKE 61
Db 471 LQSSISNLQAAKEELEQASQAHGARLTAQVASTLSELATLNLATIQQQDELAKLQQAKE 530
Qy 62 KNAELQMTLKLKEES---LTKRASQYSGQLKVLIAENTMLTSKLKEKQDEKIL---EABI 115
Db 531 KOAQLAQTLQOQEQASQGLRHQVEQLSSSL-----KQEQQLKEVAEKQEAETR 578
Qy 116 ESHHPRLASAVQDHDQIVTSRKSQBPAPHIAGDACLQRQNVVDVSTIYNNEVLHQPISLSE 175
Db 579 QDHAQQLATAAEERASLRER-----DAALKQLEALEKEKAA-KLEILOQQLQV 626
Qy 176 AQRKSKSLKINLNYAGDALRENTLVS-----EHAQRDQRETQCMKEAHMYQNE 225
Db 627 ANEARDSAQTSVT---QAQREKAELSRRKVEELQACVETARQEQHEAQVABELEQLRSE 683
Qy 226 QDNVNHKTEQESLDQKLFOQSKNMWLOQQLVHAHKADNKSKITIDHFLERKMQHHL 285
Db 684 Q---QKATE-KERVAQEKDQEQQLQALKESL-----KVTKSLSEBEKRAADA 728

Qy 286 LKEKNEEIPNHNHKNRIYQYEKEKAE 315
Db 729 LEEQQRCSISELKAETRSIVEQHKRERKLE 758

RESULT 11
A33977
myosin heavy chain, nonmuscle - chicken
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Gallus gallus (chicken)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: A33977; S06116; A43422
R:Shohet, R.V.; Conti, M.A.; Kawamoto, S.; Preston, Y.A.; Brill, D.A.; Adelstein, R.S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7726-7730, 1989
A:Title: Cloning of the cDNA encoding the myosin heavy chain of a vertebrate cellular myo-

Db 1180 AKTHEAQIQEMRQKSHQA...EELA 1202

Qy 175 EAQRKSGLNINLYAGDALRENTLVSEHAQ-----RDRETQMKEAEHMYQNE 225
Db 1203 EQLEQTKVKNANLEKA-----KQTLENRGELANEVKVLGGDRSEHKKKVEAQLDEL 1257

Qy 226 QDNVNKHTEOESLDQKLFOLOSKNMMLQQOLVIAHHKKADNK-SKITIDIHFLERMQ-- 282

Db 1258 QVKFNEGERTVELADKVTKLQE-----LDNVITGILLSQSDSKSSKLTDFSALESQLODT 1313

Qy 283 HHLLKEKNEEIFNNYHNHLKNRIYOYEKEK 311
Db 1314 QELLOENROKLSLSTKLK----QVEDEK 1338

RESULT 13

Tl8372 repeat organellar protein - Plasmodium chabaudi

C:Species: Plasmodium chabaudi

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: Tl8372

R:Werner, E.B.; Taylor, W.R.; Holder, A.A.
Mol. Biochem. Parasitol. 94, 185-196, 1998

A>Title: A Plasmodium chabaudi protein contains a repetitive region with a predicted spe-

A:Reference number: Z18922; MUID:98418765; PMID:9747969

A:Accession: Tl8372

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1939 <NER>

A:Cross-references: UNIPROT:Q25662; EMBL:U43145; NID:g1151157; PID:g1151158; PIDD:AAC634

Query Match 10.3%; Score 168.5; DB 2; Length 1939;
Best Local Similarity 22.5%; Pred. No. 0.081;
Matches 79; Conservative 69; Mismatches 128; Indels 75; Gaps 15;

Qy 6 LOCEVSHTHRENLYLHNCMLCKEIAMLKLEIATLKHOYQ-----EKENYPEDIKI 58

Db 170 IESEYTEIKNEKEKLANEVTVTIKMSLDKLTCEVQEQKDNLKINKVKVIKKNRLRELKF 229

Qy 59 LKEKN---AELOWTLKLESLTRASYSQQLKVLIAENTMTLSKLEKQDK-EILEAE 114

Db 230 MKEKNEIIIESLDGTINDKKNAYEKLETSFEKRKMI-----EMLDSKLIKEENFANKQAK 285

Qy 115 IESHHPPLASAVODHQIVTSRKQPAPFIAGDACLORKMNV---DVSST-----IYN 166

Db 286 LEKENEIIIKGLDIRESKDPFKSBEKF-----ASMENILTKSLDKSNACOMEEYKL 340

Qy 167 EV--LHOPLESEAQR-----KSKSLNINLYAGDALRENTLVSEHAQRDQ----- 208

Db 341 EI KDLSGLVEKERIEIFEIKNEYDDKINNMEKLESSINDKGIDNTVL--HSSEEKINKUL 398

Qy 209 RETOCQMKEAHEHYQNEODNV-NKHTEOESLDQKLFOLOSKNMMLQQOLVIAHKKAD-- 265

Db 399 KEKETELNEIHKNYLNLETITKNEIEXEBELEK-----NKKATHV 439

Qy 266 NKSKTIDIHFLEKGMQHLLKEKNEEIFNNYHNHLKNRIYOYEKEKAETEN 316

Db 440 EVTNLTKLEIKLEKTDEAKEGHKNE-----LNELNQSLKUNKERNKIKN 485

RESULT 14

S21801 myosin heavy chain, neuronal [similarity] - rat

N:Alternate names: myosin II

N:Contains: myosin ATPase (EC 3.6.4.1)

C:Species: Rattus norvegicus (Norway rat)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

C:Accession: S21801; PNO013; S18134

J:Sun, W.; Chantier, P.D.

R. Mol. Biol. 224, 1185-1193, 1992

A>Title: Cloning of the cDNA encoding a neuronal myosin heavy chain from mammalian brain

A:Reference number: S21801; MUID:92235856; PMID:1569576

A:Accession: S21801
A:Molecule type: mRNA
A:Residues: 1-1999 <SUN>
A:Cross-references: UNIPROT:Q63731; EMBL:X62659
R:Sun, W.; Chantler, P.D.
Biochem. Biophys. Res. Commun. 175, 244-249, 1991
A>Title: A unique cellular myosin II exhibiting differential expression in the
A:Reference number: FN0013; MUID:91151356; PMID:1998509
A:Accession: FN0013
A:Molecule type: mRNA
A:Residues: 1914-1998,'I' <SU2>
A:Experimental source: brain
C:Superfamily: myosin heavy chain; myosin motor domain homology
K:Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid;
F:84-763/Domain: myosin motor domain homology <MOT>
F:174-181/Region: nucleotide-binding motif A (P-loop)
F:541-575/Region: actin binding #status predicted
F:653-675/Region: actin binding #status predicted
F:836-1999/Domain: coiled coil #status predicted <COI>
F:836-1276/Region: S2
F:1277-1999/Region: light meromyosin
F:125/Modified site: N6,N6-trimethyllysine (Lys) #status predicted
E:180/Binding site: ATP (Lys) #status predicted
F:693,703/Active site: Cys #status predicted
F:1916/Binding site: phosphate (Ser) (covalent) #status predicted
F:1943/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 10.2%; Score 167; DB 1; Length 1999;
Best Local Similarity 22.5%; Pred. No. 0.1;
Matches 75; Conservative 55; Mismatches 101; Indels 102; Gaps 11;

Qy 27 LKKEIAMLKLEIATLKHQYGEKNKYFDIKILKKNVAELQMTLK-----LKEE 75
Db 1059 LSDQIAEQIAELKMLAKKEELQAALARVEEAQAQKMKALKITRELSQISELQD 1118
Qy 76 SLTKRASQYSG-----QLKVLT-----IAENTMTLSKLKEKDKEI-----LEAE 114
Db 1119 LESGRASRNKAQKQRDLGSEALKTLEDLDTDSTAQQELSKRKQEYNILKKTLEE 1178
Qy 115 IESHPHPLASAGVDHDQIVTSRSQSPEAFHAGDACLQRKGNVDSVTIYNNEVLHQPLS 174
Db 1179 AKTHEAQIQMRQXHSAV-----BELAEQLE 1205
Qy 175 EAQRK----SKSLKINLYAGDALRENTLV-----SEHAQR-----DORTQCQMEAEH 220
Db 1206 QTRKEVKANLEKAKQTLENERGELANEVKVLQGGRDSEHKRKKVYAQLQELQVKFNEMER 1265
Qy 221 MYQNEQNVNKHTEQDESLODKLFOLQSKNMWLLQOQLVHAHKADNKSKTIDIHFLERK 280
Db 1266 RVTELADKVTKLQVELDNVTVGLLSQSDSKS-----SKLTQKFSALESQ 1308
Qy 281 MQ--HHLLKEKNEEIFNNHNHLNRIRYQYEKEK 311
Db 1309 LDQTQELLQSENROKJLSLTKLK-----QVEDEK 1337

RESULT 15
T38435
coiled coil protein - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38435
R:McDonagall, R.; Wood, V.; Bartell, B.G.; Rajadream, M.A.
submitted to the EMBL Data Library, December 1997
A:Reference number: Z21793
A:Accession: T38435
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-750 <MCD>
A:Cross-references: UNIPROT:O42657; EMBL:AL009227; PIDN:CAA15821.1; GSPDB:GN000
A:Experimental source: strain 972n-; cosmid c27d7
C:Genetics:
A:Gene: SPDB:SPAC27D7.02c

A: title: Cloning of the cDNA encoding a neuronal myosin heavy chain from mammalian brain
J. MOL. BIOL. 224, 1185-1193, 1992
A: reference number: S21801; MUID: 92235856; PMID: 1569576

A:Gene: SPDB:SPAC27D7.02c
C:Genetics:
Experimental source: Strain 3728 / COSMAC 327,
Experimental source: Strain 3728 / COSMAC 327

A;Map position: 1

```
Query Match      10.2%; Score 166; DB 2; Length 750;
Best Local Similarity 23.1%; Pred.No. 0.037;
Matches 80; Conservative 77; Mismatches 122; Indels 68; Gaps 17;

Qy 3 TRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIA-----TLKHQYQEKENKYFED 55
Db 428 SRLLQSRTOQTGLDSYIT-SNSQLKDEITSLKQTVSESEAEKRLFPSSAQEKQLQMKET 486

Qy 56 IK---ILKEKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLK-----E 104
Db 487 VNKLTSLQEQNNPEFDRLKEQEDLQNKEBELT-ELRKLRLREQTQDSQKRLLLVEQLELE 545

Qy 105 KQDEILEAEIESHHPLASAVQPHDQIVTSRKQEPAPHIAGDACL--QRKNVDVYST 162
Db 546 RQD---LKQAGENHYSNLSS---DYE---TQIKSLESSLTNSQAECVSPQEKIN-ELNSQ 595

Qy 163 IYNNEVLHQPLSEAQKSKSLKI---NLNVAGDALRENTLVSEHAQRDQRETQCMKEAE 219
Db 596 I---DELKLLNEANKKYQELAISFENSNNKTVQSVPEPDNGLSLEALKNENQT----- 644

Qy 220 HMYQNEQDNVNHKTEQQESLDQKLFQLQSKNMMLQQQLVHAHKADNKSKITIDIH---- 275
Db 645 -LLKNLEDSTARVEHLQKSFKNVFNQLRK-----QPSNHNSSSVSRSSSVSVNSKHP 698

Qy 276 -----FLERKMOHLL-----KEKNEEINNNHKNRIYQYEKEK 311
Db 699 GSDDMLIDKEYTRNLPQFLQDRDRRPEIVNLJLSILD-LSEEQKQK 744
```

Search completed: August 1, 2005, 22:23:16
Job time : 22.5308 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:24:58 ; Search time 4799.32 Seconds
(without alignments)
3200.517 Million cell updates/sec

Title: US-09-489-079-25

Perfect score: 1632

Sequence: 1 MGTRALQCVSHTHENYL.....NHLKNRIYQYKEKAETENS 317

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09489079/runat_01082005_141250_13297/app_query.fasta_1.1635
-DB=GenEmbl -QFWT=fastcap -SUFFIX=rge -MINMATCH=0.1 -DOOPCL=0 -DOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN_1_1.8858 @runat_01082005_141250_13297 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sw.*
- 13: gb.un.*
- 14: gb.vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1632	100.0	1665	6 AR280672	AR280672 Sequence
2	1632	100.0	1665	6 AR283168	AR283168 Sequence
3	1632	100.0	1665	6 AR341941	AR341941 Sequence
4	1632	100.0	1665	6 AR343936	AR343936 Sequence

5	1632	100.0	1665	6 AR351137	AR351137 Sequence
6	1632	100.0	1665	6 AR352903	AR352903 Sequence
7	1632	100.0	1665	6 AR453717	AR453717 Sequence
8	1632	100.0	1665	6 AR561305	AR561305 Sequence
9	1632	100.0	1665	6 AX282966	AX282966 Sequence
10	1632	100.0	1665	6 AX302858	AX302858 Sequence
11	1632	100.0	1665	6 AR283456	AR283456 Sequence
12	1632	100.0	2307	6 AR283458	AR283458 Sequence
13	1632	100.0	2307	6 AR344224	AR344224 Sequence
14	1632	100.0	2307	6 AR351425	AR351425 Sequence
15	1632	100.0	2307	6 AR454005	AR454005 Sequence
16	1632	100.0	2307	6 AR561593	AR561593 Sequence
17	1632	100.0	2307	6 AX282975	AX282975 Sequence
18	1632	100.0	2307	6 AX303148	AX303148 Sequence
19	1632	100.0	2683	6 BD271311	BD271311 Reagents
20	1632	100.0	2683	6 BD271312	BD271312 Reagents
21	1632	100.0	3681	6 AR283451	AR283451 Sequence
22	1632	100.0	3681	6 AR344219	AR344219 Sequence
23	1632	100.0	3681	6 AR351420	AR351420 Sequence
24	1632	100.0	3681	6 AR454000	AR454000 Sequence
25	1632	100.0	3681	6 AR561588	AR561588 Sequence
26	1632	100.0	3681	6 AX282970	AX282970 Sequence
27	1622	99.4	972	6 AR303143	AR303143 Sequence
28	1622	99.4	972	6 AR283463	AR283463 Sequence
29	1622	99.4	972	6 AR351432	AR351432 Sequence
30	1622	99.4	972	6 AR454012	AR454012 Sequence
31	1622	99.4	1206	6 AX303162	AX303162 Sequence
32	1622	99.4	1206	6 AR280670	AR280670 Sequence
33	1622	99.4	1206	6 AR283166	AR283166 Sequence
34	1622	99.4	1206	6 AR341939	AR341939 Sequence
35	1622	99.4	1206	6 AR343934	AR343934 Sequence
36	1622	99.4	1206	6 AR351135	AR351135 Sequence
37	1622	99.4	1206	6 AR352901	AR352901 Sequence
38	1622	99.4	1206	6 AR453715	AR453715 Sequence
39	1622	99.4	1206	6 AR561303	AR561303 Sequence
40	1622	99.4	1206	6 AX282964	AX282964 Sequence
41	1622	99.4	1233	6 AX302855	AX302855 Sequence
42	1622	99.4	1233	6 AR283469	AR283469 Sequence
43	1622	99.4	1233	6 AR454018	AR454018 Sequence
44	1586	97.2	1681	6 AX303172	AX303172 Sequence
45	1586	97.2	1681	6 AR280673	AR280673 Sequence
				6 AR283169	AR283169 Sequence

ALIGNMENTS

RESULT 1
AR280672
LOCUS AR280672 1665 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 178 from patent US 6518237.
ACCESSION AR280672
VERSION AR280672.1 GI:29716142
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1665)
AUTHORS Yuglu,J., Dillon,D.C., Mitcham,J.L. and Xu,J.
TITLE Compositions for treatment and diagnosis of breast cancer and methods for their use
JOURNAL Patent: US 6518237-A 178 11-FEB-2003;
FEATURES
source Location/Qualifiers
1..1665
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 9.68e-116 Length: 1665
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```
US-09-489-079-25 (1-317) x AR280672 (1-1665)
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAenGluAenTyrLeu 20
Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTCTCACACTCATGAAATGAAATATATCTC 518
Qy 21 LeuHisGluAenCysMetLeuLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
Db 519 TTACATGAAATTCGATGTTGAAAAGGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 578
Qy 41 LeuLysHisGlnTyrGlnGluLysGluAenLysTyrPheGluAspLysLeuLys 60
Db 579 CTGAAACACCAATACCAAGGAAAGGAAATGAAATGCTGCTGCTGCTGCTGCTGCTGCT 638
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuGluLeuAlaThr 80
Db 639 GAAAGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAenThrMetLeuThrSer 100
Db 699 GCATCTCAATATAGTGGGAGCTTAAAGTTCTGATAGCTGAGAACACATGCTCCTCTCT 758
Qy 101 LysLeuLysGluLysGlnAspLysGluLeuGluAlaGluLeuGluSerHisPro 120
Db 759 AAATTTGAAGGAAACCAAGCAAGAAATGCTGAGGCGAGAAATGGAATGCACTCCT 818
Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTTCTGCTGTACAAAGACCATGATCAATTTGTGACATCAAGAAATGCTCAAG 878
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 879 CCTGCTTCCACATTCGAGGAGATGCTTTGTTGCAAGGAAATGGAATGCTGCTGCT 938
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTACGATATATACAAATGAGTCTCCATCAACCACTTTCTGAAGCTCAAGGAAATCC 998
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAenThrLeuVal 200
Db 999 AAAAGCCTTAAATATATCTCAATTTATGCGGAGATGCTCTTAAGAGAAATACATCTG 1058
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGACATGCACAAAGAGACCAACCGTGAACACAGTGTCAAATGAAGGAGCTGAACAC 1118
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnSerLeuAsp 240
Db 1119 ATGTATCAAAACGAACAAGATAATGTGAACAAACACACTGACAGCAGGAGTCTCTAG 1178
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnLeuValHisAla 260
Db 1179 CAGAAATTTATTTCAACTACAAAGCAAAATATGTTGGCTTCAACAGCAATTTAGTTCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysLysThrIleAspLysPheLeuGluArgLys 280
Db 1239 CATAGAAAGAGCTGACAAACAAAGCAAGATACAAATGATATTCATTTCTTTGAGAGGAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTTAAAGAGAAATGAGGAGATATTTAATTACAATAACCAATTA 1358
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAenSer 317
Db 1359 AAAAACCGTATATATCAATATGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
RESULT 2
AR283168 LOCUS AR283168 1665 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 178 from patent US 6528054.
ACCESSION AR283168
VERSION AR283168.1 GI:29719995
```

```
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1. (bases 1 to 1665)
AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and Hepler, W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6528054-A 178 04-WAR-2003;
FEATURES Location/Qualifiers
source 1.1665
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 9.68e-116 Length: 1665
Pred. No.: 1632.00 Matches: 317
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-489-079-25 (1-317) x AR283168 (1-1665)
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAenGluAenTyrLeu 20
Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTCTCACACTCATGAAATGAAATATATCTC 518
Qy 21 LeuHisGluAenCysMetLeuLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
Db 519 TTACATGAAATTCGATGTTGAAAAGGAATTTGCCATGCTTAAACTGGAATAGCCACA 578
Qy 41 LeuLysHisGlnTyrGlnGluLysGluAenLysTyrPheGluAspLysLeuLys 60
Db 579 CTGAAACACCAATACCAAGGAAAGGAAATGAAATGCTGCTGCTGCTGCTGCTGCTGCT 638
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuGluLeuAlaThrLysArg 80
Db 639 GAAAGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAenThrMetLeuThrSer 100
Db 699 GCATCTCAATATAGTGGGAGCTTAAAGTTCTGATAGCTGAGAACACATGCTCCTCTCT 758
Qy 101 LysLeuLysGluLysGlnAspLysGluLeuGluAlaGluLeuGluSerHisPro 120
Db 759 AAATTTGAAGGAAACCAAGCAAGAAATGCTGAGGCGAGAAATGGAATGCACTCCT 818
Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTTCTGCTGTACAAAGACCATGATCAATTTGTGACATCAAGAAATGCTCAAG 878
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 879 CCTGCTTCCACATTCGAGGAGATGCTTTGTTGCAAGGAAATGGAATGCTGCTGCT 938
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTACGATATATACAAATGAGTCTCCATCAACCACTTTCTGAAGCTCAAGGAAATCC 998
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAenThrLeuVal 200
Db 999 AAAAGCCTTAAATATATCTCAATTTATGCGGAGATGCTCTTAAGAGAAATACATCTG 1058
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGACATGCACAAAGAGACCAACCGTGAACACAGTGTCAAATGAAGGAGCTGAACAC 1118
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnSerLeuAsp 240
Db 1119 ATGTATCAAAACGAACAAGATAATGTGAACAAACACACTGACAGCAGGAGTCTCTAG 1178
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnLeuValHisAla 260
Db 1179 CAGAAATTTATTTCAACTACAAAGCAAAATATGTTGGCTTCAACAGCAATTTAGTTCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysLysThrIleAspLysPheLeuGluArgLys 280
Db 1239 CATAGAAAGAGCTGACAAACAAAGCAAGATACAAATGATATTCATTTCTTTGAGAGGAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTTAAAGAGAAATGAGGAGATATTTAATTACAATAACCAATTA 1358
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAenSer 317
Db 1359 AAAAACCGTATATATCAATATGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
```

QY 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrrPLeuGlnGlnLeuValHisAla 260
 DB 1179 CAGAAATTTTCAACTACAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1238
 QY 261 HisLysLysAlaAspAsnLysSerLysLysLeuThrIleAspIleHisPheLeuGluArgLys 280
 DB 1239 CATAAGAAAGCTGACACAAACAAAGCAAGATAACAATTTGATATTATTCATTTCTTGAGAGGAAA 1298
 QY 281 MetGlnHisLeuLeuLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 300
 DB 1299 ATGCACATCATCTCTTAAAGAGAAAATGAGGAGATATTTAATACATTAACCATTTTA 1358
 QY 301 LysAsnArgIleTyrGlnTyrGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 317
 DB 1359 AAAACCGGTATATATCAATATGAAAGAGAAAAGCAAGCAAGAAACAACTCA 1409
 RESULT 3
 AR341941
 LOCUS AR341941 1665 bp DNA linear PAT 17-AUG-2003
 DEFINITION Sequence 178 from patent US 6573368.
 ACCESSION AR341941
 VERSION AR341941.1 GI:33736518
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1665)
 AUTHORS Yuqiu,J., Dillon,D.C., Mitcham,J.L. and Xu,J.
 TITLE Compositions for the treatment and diagnosis of breast cancer and methods for their use
 JOURNAL Patent: US 6573368-A 178 03-JUN-2003;
 FEATURES
 source Location/Qualifiers
 1..1665
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 9 68e-116 Length: 1665
 Score: 1632.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-489-079-25 (1-317) x AR341941 (1-1665)
 QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
 DB 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 518
 QY 21 LeuHisGluAsnCysMetLeuLysLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 40
 DB 519 TTACATGAAATTTGCATGTTTGAAGAAAGGAAATGCGCATCTTAAATCTGAAATAGCCACA 578
 QY 41 LeuLysHisGlnTyrGlnGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 60
 DB 579 CTGAAACACCAATACCGAAGAAAGGAAATTAATATCTTTGAGGACATTAAGATTATAAA 638
 QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
 DB 639 GAAAGAAATGCTGAATCTCAGATGACCTTAAACCTGAAAGAGGAAATCAATTAACATAAGG 698
 QY 81 AlaSerGlnTyrSerGlyGlnLysValLeuLysValLeuLysAlaGluAsnThrMetLeuThrSer 100
 DB 699 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGAACCAATGCTCACTTCT 758
 QY 101 LysLeuLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 120
 DB 759 AAATTGAAGGAAACCAAGCAAGAAATCTAGAGCGCAAAATTAATGAAATCAACACCTCT 818
 QY 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140

DB 819 AGACTGGCTTCTGCTGTACAAAGACCATGATCAATTTTGACATCAAGAAAAGTCAAGAA 878
 QY 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
 DB 879 CTGCTTTCCACATTTGCAGGAGATGCTTTGTTTGCACAAAGAAAATGAATGTTGATGTGAGT 938
 QY 161 SerThrIleTyrAsnAsnGlnValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
 DB 939 AGTACGATATATAACAATGAGGTGCTCCATCAACCACTTTCTGAAGCTCAAGAGAAATCC 998
 QY 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
 DB 999 AAAAGCCCTAAATAATTAATCTCAATTTATGCCGAGATGCTCTAAGAGAAAATACATTTGGTT 1058
 QY 201 SerGluHisAlaGlnArgAspGlnArgLysGlnThrGlnCysGlnMetLysGlnAlaGluHis 220
 DB 1059 TCAGAACATGCACAAAGAGACCAACGTGAACACACAGTGTCAATGAAGAGCTGAACAC 1118
 QY 221 MetTyrGlnAsnGlnGlnAspAsnValAsnLysHisThrGluGlnGlnGlnSerLeuAsp 240
 DB 1119 ATGTATCAAAACGAAACAAGATAATGTGAACAAACACACTGAACAGAGGAGTCTTAGAT 1178
 QY 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrrPLeuGlnGlnGlnLeuValHisAla 260
 DB 1179 CAGAAATTTTCAACTACAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1238
 QY 261 HisLysLysAlaAspAsnLysSerLysLysLeuThrIleAspIleHisPheLeuGluArgLys 280
 DB 1239 CATAAGAAAGCTGACACAAACAAAGCAAGATAACAATTTGATATTATTCATTTCTTGAGAGGAAA 1298
 QY 281 MetGlnHisLeuLeuLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 300
 DB 1299 ATGCACATCATCTCTTAAAGAGAAAATGAGGAGATATTTAATTAACCAATTAACCATTTTA 1358
 QY 301 LysAsnArgIleTyrGlnTyrGlnLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 317
 DB 1359 AAAACCGGTATATATCAATATGAAAGAGAAAAGCAAGCAAGAAACAACTCA 1409
 RESULT 4
 AR343936 1665 bp DNA linear PAT 17-AUG-2003
 LOCUS AR343936
 DEFINITION Sequence 178 from patent US 6579973.
 ACCESSION AR343936
 VERSION AR343936.1 GI:33739836
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1665)
 AUTHORS Yuqiu,J., Dillon,D.C., Mitcham,J.L., Xu,J. and Harlocker,S.L.
 TITLE Compositions for the treatment and diagnosis of breast cancer and methods for their use
 JOURNAL Patent: US 6579973-A 178 17-JUN-2003;
 FEATURES
 source Location/Qualifiers
 1..1665
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 9 68e-116 Length: 1665
 Score: 1632.00 Matches: 317
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-489-079-25 (1-317) x AR343936 (1-1665)
 QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
 DB 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 518

Qy	21	LeuHisGluAsnCysMetLeuLeuLysGluIleAlaMetLeuLysLeuGluIleAlaThr	40
Db	519	TTATCATGAAATTCGATGTTGAAAAGAAATGCCATGCTTAAACCTGGAATAGCCACA	578
Qy	41	LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys	60
Db	579	CTGAAACACCAATACCAGGAAAGGAAATAAATACCTTTGAGGACATTAAGATTTTAAA	638
Qy	61	GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysGluSerLeuThrLysArg	80
Db	639	GAAGAAGAAATGCTGAACCTTACAGTACCTTAAACCTGAAGAGGAATCATTAACCTAAAAGG	698
Qy	81	AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer	100
Db	699	GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGAACCAATGCTCCTCCTCT	758
Qy	101	LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro	120
Db	759	AAATTTGAAGGAAACCAAGCAAGAAATACCTAGAGGCAGAAATTTGAATCACACCATCCT	818
Qy	121	ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu	140
Db	819	AGACTGGCTTCTGCTGTACAGACCATGATCAAAATTTGTGACATCAAGAAAAGTCAAGAA	878
Qy	141	ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer	160
Db	879	CCTGCTTCCCATTCGAGGAGATGCTTTGTTCAAGAAATAATGATTTGATGTGAGT	938
Qy	161	SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer	180
Db	939	AGTACGATATATAACATGAGTGTCTCATCAACCATCTTCTGAAGCTCAAAGGAAATCC	998
Qy	181	LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgLysGluAsnThrLeuVal	200
Db	999	AAAAGCCCTAAAATTAATCTCAATTAATGCGAGATGCTCTAAGAGAAAATACATTTGGTT	1058
Qy	201	SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis	220
Db	1059	TCGAACATGCAACAAAGAGACCAACGTTGAAACACAGTGTCAATGAAAGGAACTGAACAC	1118
Qy	221	MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp	240
Db	1119	ATGTATCAAAACGAACCAAGATAATGTTGTAACAAACACACACTGAACGAGGAGTCTCTAGAT	1178
Qy	241	GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla	260
Db	1179	CAGAAATTAATTTCAACTCAACAAAGCAAAATATGTGGCTTCAACAGCAATAGTTTCATGCA	1238
Qy	261	HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys	280
Db	1239	CATAGAAAGCTGACAAACAAAGCAAGATAACAATTTGATATTCATTTTCTTGAGAGGAA	1298
Qy	281	MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu	300
Db	1299	ATGCAACATCATCTCTTAAAGAGAAATATGAGGAGATATTTAATTAACAAATTA	1358
Qy	301	LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer	317
Db	1359	AAAAACCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAACTCA	1409
RESULT 5			
LOCUS	AR351137	1665 bp	DNA linear PAT 17-AUG-2003
DEFINITION	Sequence 178 from patent US 6586572.		
ACCESSION	AR351137		
VERSION	AR351137.1	GI:33752816	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1665)		
AUTHORS	Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J., Harlocker, S.L. and		

TITLE	Hepler, W.T. Compositions and methods for the therapy and diagnosis of breast cancer		
JOURNAL	Patent: US 6586572-A 178 01-JUL-2003;		
FEATURES	location/Qualifiers		
source	1. 1665		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:		Length:	1665
Pred. No.:	9.68e-116	Matches:	317
Score:	1632.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	6		
US-09-489-079-25 (1-317) x AR351137 (1-1665)			
Qy	1	MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu	20
Db	459	ATGGGAAACAAGAGCTGCGAGTGTGAGGTTTCTCACACTCATGAAAATGAAAATATATCTC	518
Qy	21	LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr	40
Db	519	TTATCATGAAATTTGCATGTTTGAAGAAAGAAATTTGCCATGCTTAAACTGGAAATAGCCACA	578
Qy	41	LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys	60
Db	579	CTGAAACACCAATACCAGGAAAGGAAATAAATACTTTGAGGACATTAAGATTTTAAAA	638
Qy	61	GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysGluSerLeuThrLysArg	80
Db	639	GAAGAAGAAATGCTGAACCTTACAGTACCTTAAACTGAAGAGGAATCATTAACCTAAAAGG	698
Qy	81	AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer	100
Db	699	GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGAACACAATGCTCCTCCTCT	758
Qy	101	LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro	120
Db	759	AAATTTGAAGGAAACCAAGCAAGAAATATCTAGAGGCAGAAATTTGAATCACACCATCCT	818
Qy	121	ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu	140
Db	819	AGACTGGCTTCTGCTGTACAGACCATGATCAAAATTTGTGACATCAAGAAAAGTCAAGAA	878
Qy	141	ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer	160
Db	879	CCTGCTTCCCATTCGAGGAGATGCTTTTTCGAAAGAAAATGAAATGTTGATGTGAGT	938
Qy	161	SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer	180
Db	939	AGTACGATATATAACATGAGTGTCTCATCAACCATCTTCTGAAGCTCAAAGGAAATCC	998
Qy	181	LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgLysGluAsnThrLeuVal	200
Db	999	AAAAGCCCTAAAATTAATCTCAATTAATGCGAGATGCTCTAAGAGAAAATACATTTGGTT	1058
Qy	201	SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis	220
Db	1059	TCGAACATGCAACAAAGAGACCAACGTTGAAACACAGTGTCAATGAAAGGAACTGAACAC	1118
Qy	221	MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp	240
Db	1119	ATGTATCAAAACGAACCAAGATAATGTTGTAACAAACACACACTGAACGAGGAGTCTCTAGAT	1178
Qy	241	GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla	260
Db	1179	CAGAAATTAATTTCAACTCAACAAAGCAAAATATGTGGCTTCAACAGCAATAGTTTCATGCA	1238
Qy	261	HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys	280
Db	1239	CATAGAAAGCTGACAAACAAAGCAAGATAACAATTTGATATTCATTTTCTTGAGAGGAA	1298
Qy	281	MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu	300
Db	1299	ATGCAACATCATCTCTTAAAGAGAAATATGAGGAGATATTTAATTAACAAATTA	1358
Qy	301	LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer	317
Db	1359	AAAAACCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAACTCA	1409

```
Db 1239 CATAGAAAGCTGCAACAACAAAGCAAGATAACAATTTGATATTCATTTCTTGAGAGGAAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluLeuPheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTCTAAAGAGAAAATGAGGAGATATTTAATACATAACCATTTA 1358
Qy 301 LysAsnArgIleTyrGlnTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1359 AAAAACCGTATATATCAATATGAAAAGAGAAAGCAGAAAACAGAAAACCTCA 1409

RESULT 6
AR352903
LOCUS AR352903 1665 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 178 from patent US 6590076.
ACCESSION AR352903
VERSION AR352903.1 GI:33758308
KEYWORDS
SOURCE
ORGANISM
Unclasseified.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Yudin,J., Dillon,D.C., Mitcham,J.L. and Xu,J.
TITLE Compositions for the treatment and diagnosis of breast cancer and
methods for their use
JOURNAL Patent: US 6590076-A 178 08-JUL-2003;
FEATURES
source Location/Qualifiers
DB: 1..1665
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 9,68e-116 Length: 1665
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AR352903 (1-1665)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAAATTTATCTC 518
Qy 21 LeuHisGluAsnCysMetLeuLysGluLysGluLysAlaMetLeuLysGluLysLeuAlaThr 40
Db 519 TTACATGAAAATTCATGTTGAAAAGGAAAATTCGCATGCTAAAACCTGAAAATAGCCACA 578
Qy 41 LeuLysHisGlnTyrGlnGluLysGluLysGluLysTyrPheGluAspIleLysIleLeuLys 60
Db 579 CTGAAACACCAATACCAAGAAAAGGAAAATAAATCTTTGAGACATTAAGATTTTAAAA 638
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysGluLysThrLysArg 80
Db 639 GAAAGAGATGCTGAACTTCAGATGACCTTAAACCTGAAAGAGGAATCATTAATACTAAAGG 698
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLysAlaGluAsnThrMetLeuThrSer 100
Db 699 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTTGAGAACACATGCTCACTTCT 758
Qy 101 LysLeuLysGluLysGlnAspLysGluLysGluLysGluLysGluLysSerHisPro 120
Db 759 AAATTGAGAAAACACAGACAAAGAAAATACCTAGAGCAGAAATTTGAAATCACACCATCT 818
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTTCTGCTGTACAGACCATGATCAAAATTTGTGACATCAAGAAAAGTCAAGAA 878
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValSer 160
Db 879 CCTGCTTTCCACATTCAGAGATGCTTGTGTTGCAAGAAAATGAATGTTGATGTGAGT 938
```

```
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTACGATATATAACAATAGAGTGTCTCCATCAACCACTTTCTGAAGCTCAAGGAAATCC 998
Qy 181 LysSerLeuLysIleLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 999 AAAAGCCTTAAAAATTAATCTCAATTTATGCCGAGATGCTCTAAGAGAAAATACATGGTT 1058
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGAACATGTCACAAAGAGACCAACGTGAACACACAGTGTCAATGAAGGAAGCTGAACAC 1118
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGlnSerLeuAsp 240
Db 1119 ATGTATCAAAAAGCAACAGATAATGTGAACAAACACACTGAACACGACGAGGTCTCTAGAT 1178
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
Db 1179 CAGAAATTTATTTCACTACAAGCAAAAATATGTGGCTTCAACACAAATTAGTTCAATGCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1239 CATAGAAGAGCTGACACAAAGCAAGATAACAATTGATATTCATTTCTTGAGAGGAAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluLysLeuPheAsnTyrAsnAsnHisLeu 300
Db 1299 ATGCAACATCATCTCTCTAAAAGAGAAAATGAGGAGATATTTAATTAACAATAACCATTTA 1358
Qy 301 LysAsnArgIleTyrGlnTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1359 AAAAACCGTATATATCAATATGAAAAGAGAAAAGCAGAAAACAGAAAACCTCA 1409

RESULT 7
AR453717
LOCUS AR453717 1665 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 178 from patent US 6680197.
ACCESSION AR453717
VERSION AR453717.1 GI:42686507
KEYWORDS
SOURCE
ORGANISM
Unclasseified.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L.,
Hepler,W.T. and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: US 6680197-A 178 20-JAN-2004;
FEATURES
source Location/Qualifiers
DB: 1..1665
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 9,68e-116 Length: 1665
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AR453717 (1-1665)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAAATTTATCTC 518
Qy 21 LeuHisGluAsnCysMetLeuLysGluLysGluLysGluLysAlaMetLeuLysGluLysLeuAlaThr 40
Db 519 TTACATGAAAATTCATGTTGAAAAGGAAAATTCGCATGCTAAAACCTGAAAATAGCCACA 578
```

```
Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLeuLys 60
Db 579 CTGAACACCAATACCAGAAAGAAAGAAATAAATCTTTGAGGACATTAAGATTTTAAA 638
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
Db 639 GAAAGAAGTCTGAACTTCAGATGACCCCTAAATCTGAAAGAGGAATCATTAACCTAAAGG 698
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLysAlaGluAsnThrMetLeuThrSer 100
Db 699 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGTAGCTGAGAACACAAATGCTCACTTCT 758
Qy 101 LysLeuLysGluLysGlnAspLysGluLysGluLysGluLysGluLysGluSerHisPro 120
Db 759 AAATTTGAAGGAAACCAAGACAAAGAAATATCTAGAGGCGAGAAATGATTCACACATCCT 818
Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnLeuLysValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTTCTGTGTACAAAGACCATGATCAAAATTTGTGATCATCAAGAAAGTCAAGAA 878
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 879 CCTGCTTTCCACATTCGACAGGAGATGCTTTGTCAGAAAGAAATGAAATGTTGATGTGAGT 938
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTACGATATATACATAGAGTGTCTCATCCACATTTCTGAGGCTCAAGAGAAATCC 998
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 999 AAAAGCCTTAAATTTAATCTCAATTTATGCGGAGATGCTTAAAGAGAAATATACATGCTT 1058
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGAACATGCAACAAAGAGACCAACGCGTGAACACAGTGTCAAAATGAAAGGAGCTGAACAC 1118
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db 1119 ATGTATCAAAACGAAACAAAGATATATGTGNAACAAACACACTGACAGCAGGAGTCTTAGAT 1178
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
Db 1179 CAGAAATTTATTTCAACTACAAAGCAGAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1239 CATAGAAAGCTGACAAACAAAGCAAGATAAATGATGATATTCATTTTCTTTGAGAGGAAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluLysLeuPheAsnTyrAsnAsnHisLeu 300
Db 1359 AAAAACCGTATATATCAATATGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409

RESULT 8
AR561305 1665 bp DNA linear PAT 08-OCT-2004
LOCUS Sequence 178 from patent US 6756477.
DEFINITION AR561305
ACCESSION AR561305.1 GI:53974213
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Jiang Y., Dillon D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6756477-A 178 29-JUN-2004;
FEATURES Location/Qualifiers
1..1665 source
```

```
ORIGIN
Alignment Scores: 9.68e-116 Length: 1665
Pred. No.: 1632.00 Matches: 317
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
/organism="unknown"
/mol_type="genomic DNA"

US-09-489-079-25 (1-317) x AR561305 (1-1665)
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 459 ATGGGAACAAGAGCTCTGCGAGTGTGAGGTTCTCACACTCATGAAATGAAAATTAATCTC 518
Qy 21 LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
Db 519 TTACATGAAATTTGCATGTTGAAAAGGAAATTTGCCATGCTAAACCTGGAATAGCCACA 578
Qy 41 LeuLysHisGlnTyrGlnGlnLysGluAsnLysTyrPheGluAspIleLysIleLys 60
Db 579 CTGAAACACCAATACCAAGGAAAGAAATAATACTTTGAGGACATTAAGATTTTAAA 638
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
Db 639 GAAAGAAATGCTGAACTTCAGATGACCTTAAACTGAAAGAGGAATCATTAACCTAAAGG 698
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db 699 GCATCTCAATATAGTGGCGAGCTTAAAGTTCTGTAGTCTGAGAACACATGCTCACTTCT 758
Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro 120
Db 759 AAATTTGAAGGAAACCAAGACAAAGAAATATCTAGAGGCGAGAAATTTGAATCACACATCCT 818
Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnLeuValLeuIleValThrSerArgLysSerGlnGlu 140
Db 819 AGACTGGCTTCTGTGTACAAAGACCATGATCAAAATTTGTGATCATCAAGAAAGTCAAGAA 878
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 879 CCTGCTTTCCACATTTGACAGGAGATGCTTTGTCAGAAAGAAATGAAATGTTGATGTGAGT 938
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 939 AGTACGATATATACATAGAGTGTCTCATCCACATTTCTGAGGCTCAAGAGAAATCC 998
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 999 AAAAGCCTTAAATTTAATCTCAATTTATGCGGAGATGCTTAAAGAGAAATATACATGCTT 1058
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1059 TCAGAACATGCAACAAAGAGACCAACGCGTGAACACAGTGTCAAAATGAAAGGAGCTGAACAC 1118
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db 1119 ATGTATCAAAACGAAACAAAGATATATGTGNAACAAACACACTGACAGCAGGAGTCTTAGAT 1178
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
Db 1179 CAGAAATTTATTTCAACTACAAAGCAGAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1238
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1239 CATAGAAAGCTGACAAACAAAGCAAGATAAATGATGATATTCATTTTCTTTGAGAGGAAA 1298
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluLysLeuPheAsnTyrAsnAsnHisLeu 300
Db 1359 AAAAACCGTATATATCAATATGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1409

RESULT 8
AR561305 1665 bp DNA linear PAT 08-OCT-2004
LOCUS Sequence 178 from patent US 6756477.
DEFINITION AR561305
ACCESSION AR561305.1 GI:53974213
VERSION
KEYWORDS
SOURCE
ORGANISM
Unclassified.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Jiang Y., Dillon D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6756477-A 178 29-JUN-2004;
FEATURES Location/Qualifiers
1..1665 source
```

```
QY 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
|||||
Db 1359 AAAAACCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAACTCA 1409

RESULT 9
AX282966 9.68e-116 1665 bp DNA linear PAT 02-NOV-2001
LOCUS AX302858 1665 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 15 from Patent WO0175171.
ACCESSION AX282966
VERSION AX282966.1 GI:16609901
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Houghton,R.L., Dillon,D.C., Molesh,D.A., Xu,J., Zehentner,B. and
Persing,D.H.
TITLE Methods, compositions and kits for the detection and monitoring of
breast cancer
JOURNAL Patent: WO 0175171-A 15 11-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..1665
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.68e-116 Length: 1665
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AX282966 (1-1665)

QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
|||||
Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTCTCACACTCATGAAAGTAAATATCTTC 518

QY 21 LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
|||||
Db 519 TTACATGAAATTCATGTTGAAAAGGAAATGCGATGCTAAACTGGAATAGCCACA 578

QY 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
|||||
Db 579 CTGAAACACCAATACCAGGAAAAGGAAAATAAATACTTTGAGGACATTAAGATTTTAAA 638

QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
|||||
Db 639 GAAAAGATGCTGAACTTCAGATGACCCCTAAACACTGAAAGAGGAATCATTAATACTAAA 698

QY 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
|||||
Db 699 GCATCTCAATATAGTGGCGAGCTTAAAGTTCGTAGCTGAGACACATGCTCACTTCT 758

QY 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro 120
|||||
Db 759 AAATTGAAAGAAAACAACAGACAAAGAAATACTAGAGGCGAGAAATTAATCATCACCATCT 818

QY 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
|||||
Db 819 AGACTGGCTTCTGCTGTACAGACCATGATCAATTTGTCACATCAAGAAAAGTCAAGAA 878

QY 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
|||||
Db 879 CTGCTTTTCCATTCAGAGAGATGCTGTTTGCAAAAGAAAATGATGATGATGAGT 938

QY 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
|||||
Db 939 AGTACGATATATAACAATGAGTGTCTCATCACACATTTCTGAGCTCAAGGAATTC 998

181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
|||||
Db 999 AABAGCCTAAAAATTAATCTCAATTTATGCCGAGATGCTCTAAGAGAAAAATACATGCTT 1058

QY 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
|||||
Db 1059 TCAGAACATGCAACAAGAGAGACCAACGTCGAAAACACAGTGTCAATGAAGGAAGCTGAACAC 1118

QY 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
|||||
Db 1119 ATGATATCAAAACGAACAAGATAATGTGAACAACACACTGAACAGCAGAGTCTTAGAT 1178

QY 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
|||||
Db 1179 CAGAAATTTATTTCAACTACAAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1238

QY 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
|||||
Db 1239 CATAAGAAAGCTGACACAAAGCAAGATAACAATTGATATTTCAATTTCTTGAGAGGAAA 1298

QY 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
|||||
Db 1299 ATGCAACATCATCTCTTAAAGAGAAAANTGAGGAGATATTTAATACATAACCATTTA 1358

QY 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
|||||
Db 1359 AAAAACCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAACTCA 1409

RESULT 10
AX302858 1665 bp DNA linear PAT 30-NOV-2001
LOCUS AX302858 1665 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 178 from Patent WO0179286.
ACCESSION AX302858
VERSION AX302858.1 GI:17383358
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and
Hepler,W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast
cancer
JOURNAL Patent: WO 0179286-A 178 25-OCT-2001;
CORIXA CORPORATION (US)
FEATURES
source
1..1665
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 9.68e-116 Length: 1665
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AX302858 (1-1665)

QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
|||||
Db 459 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTCTCACACTCATGAAAGTAAATATCTTC 518

QY 21 LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40
|||||
Db 519 TTACATGAAATTCATGTTGAAAAGGAAATGCGATGCTAAACTGGAATAGCCACA 578

QY 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
|||||
Db 579 CTGAAACACCAATACCAGGAAAAGGAAAATAAATACTTTGAGGACATTAAGATTTTAAA 638

QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
|||||
Db 639 GAAAAGATGCTGAACTTCAGATGACCCCTAAACACTGAAAGAGGAATCATTAATACTAAA 698

QY 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
|||||
Db 699 GCATCTCAATATAGTGGCGAGCTTAAAGTTCGTAGCTGAGACACATGCTCACTTCT 758

QY 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisHisPro 120
|||||
Db 759 AAATTGAAAGAAAACAACAGACAAAGAAATACTAGAGGCGAGAAATTAATCATCACCATCT 818

QY 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
|||||
Db 819 AGACTGGCTTCTGCTGTACAGACCATGATCAATTTGTCACATCAAGAAAAGTCAAGAA 878

QY 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
|||||
Db 879 CTGCTTTTCCATTCAGAGAGATGCTGTTTGCAAAAGAAAATGATGATGATGAGT 938

QY 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
|||||
Db 939 AGTACGATATATAACAATGAGTGTCTCATCACACATTTCTGAGCTCAAGGAATTC 998
```



```
|||||
Db 1906 ATGCAACATCATCTCTCTTAAGAGAGAAAATGAGGAGATATTTAAATTACAATAACCAATTTA 1965
Qy 301 LysAsnArgIleTyrGlnTyrGlnLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAAAACCGTATATATCAATATGAAAAGAGAAAGACAGAAACAGAAAACCTCA 2016

RESULT 12
AR344224
LOCUS AR344224 2307 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 468 from patent US 6579973.
ACCESSION AR344224
VERSION AR344224.1 GI:33740124
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2307)
AUTHORS Yuqiu,J., Dillon,D.C., Mitcham,J.L., Xu,J., and Harlocker,S.L.
TITLE Compositions for the treatment and diagnosis of breast cancer and methods for their use
JOURNAL Patent: US 6579973-A 468 17-JUN-2003;
FEATURES Location/Qualifiers
source 1..2307
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.38e-115 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AR344224 (1-2307)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 1066 ATGGGAAACAGAGCTCTGCAGTGTGAGGTTTTCACACTCATGAAATGAAATATATCTC 1125
Qy 21 LeuHisGluAsnCysMetLeuLysGluIleAlaMetLeuLysGluIleAlaThr 40
Db 1126 TTACATGAAATTTGCATGTTGAAAAGGAAATGCCATGCTAAATCTGAAATAGCCACA 1185
Qy 41 LeuLysHisGlnTyrGlnGlnLysGluAsnLysTyrPheGluAspIleLysLeuLys 60
Db 1186 CTGAAACACCAATACCCAGAAAAGGAAAATAAATCTTTTGAGGACATTAAGATTTTAAA 1245
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysSerLeuThrLysArg 80
Db 1246 GAAAGAAATCTGAATCTTCAGATGACCTTAAACTGAAAGAGGAATCATTAATAAAGG 1305
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db 1306 GCATCTCATATATAGTGGCGAGCTTAAAGTTCTGATAGCTGAGAACACAATGCTCACTTC 1365
Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro 120
Db 1366 AAATTTGAAGGAAAAACAGACAAAGAAATACCTAGAGGCGAAATTTGAATCAACCATCTC 1425
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAGACCATGATCAATTTGTGACATCAGAAAAGAGTCAAGAA 1485
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CCTGCTTCTCCACATTCAGGAGATGCTTGTGTTGCAAGAAAATGAATGTTGATGTGAGT 1545
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTACCATATATAACAAATGAGGTGCTCCATCAACCACTTTCTGAAGCTCAAAAGGAATCC 1605
```

```
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 1606 AAAAGCCTTAAAAATTAATCTCAATTTATGCAGGAGATGCTCTAAGAGAAAAATACATGGTT 1665
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1666 TCAGAACATGCAACAGAGACCAAGCTGAAAACACAGTGTCTCAATGAGAGAGCTGACAC 1725
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnSerLeuAsp 240
Db 1726 ATGTATCAAAACGAAACAGATAATGTGAACAAACACACTGAACAGCAGAGTCTCTAGAT 1785
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnLeuValHisAla 260
Db 1786 CAGAAATTTATTTCACTACAAGCAAAAATATGTGGCTTCAACAGCAATTAGTTCATGCA 1845
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1846 CATAAGAAAGCTGACACAAAGCAAGATAACAAATGATATTCATTTCTTGAGAGAGAAA 1905
Qy 281 MetGlnHisHisLeuLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db 1906 ATGCAACATCATCTCTCTAAAGAGAAAAATGAGGAGATATTTAAATTACAATAACCATTTA 1965
Qy 301 LysAsnArgIleTyrGlnTyrGlnLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAAACCCGTATATATCAATATGAAAAGAGAAAGCAGAAACAGAAAACCTCA 2016

RESULT 13
AR351425
LOCUS AR351425 2307 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 468 from patent US 6586572.
ACCESSION AR351425
VERSION AR351425.1 GI:33753104
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2307)
AUTHORS Jiang,Y., Dillon,D.C., Mitcham,J.L., Xu,J., Harlocker,S.L. and Hepler,W.T.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6586572-A 468 01-JUL-2003;
FEATURES Location/Qualifiers
source 1..2307
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.38e-115 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AR351425 (1-2307)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 1066 ATGGGAAACAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAATGAAATATATCTC 1125
Qy 21 LeuHisGluAsnCysMetLeuLysGluIleAlaMetLeuLysGluIleAlaThr 40
Db 1126 TTACATGAAATTTGCATGTTGAAAAGGAAATTTGCCATGCTAAATCTGAAATAGCCACA 1185
Qy 41 LeuLysHisGlnTyrGlnGlnLysGluAsnLysTyrPheGluAspIleLysLeuLys 60
Db 1186 CTGAAACACCAATACCCAGAAAAGGAAAATAAATCTTTTGAGGACATTAAGATTTTAAA 1245
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluLysSerLeuThrLysArg 80
Db 1246 GAAAGAAATCTGAATCTTCAGATGACCTTAAACTGAAAGAGGAATCATTAATAAAGG 1305
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100
Db 1306 GCATCTCATATATAGTGGCGAGCTTAAAGTTCTGATAGCTGAGAACACAATGCTCACTTC 1365
Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro 120
Db 1366 AAATTTGAAGGAAAAACAGACAAAGAAATACCTAGAGGCGAAATTTGAATCAACCATCTC 1425
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAGACCATGATCAATTTGTGACATCAGAAAAGAGTCAAGAA 1485
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CCTGCTTCTCCACATTCAGGAGATGCTTGTGTTGCAAGAAAATGAATGTTGATGTGAGT 1545
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTACCATATATAACAAATGAGGTGCTCCATCAACCACTTTCTGAAGCTCAAAAGGAATCC 1605
```

```
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
Db 1246 GAAAGAAATGCTGAACCTTCAGATGACCTTAAACCTGAAAGAGGAATCATTTAACTAAAAGG 1305
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAenThrMetLeuThrSer 100
Db 1306 GCATCTCAATATAGTGGGACGCTTAAAGTTCTGATAGCTGAGAACACAAATGCTCAGTTCT 1365
Qy 101 LysLeuLysGluLysGlnAspLysGluLeuLeuGluAlaGluLeuSerHisHisPro 120
Db 1366 AAATTGAAGGAAACAAAGCAAAAGAAATATCTAGAGGAGAAATTTGAATCACACATCTCT 1425
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnTyrValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAGACCATGATCAATTTGTGACATCAAGAAAAGTCAAGAA 1485
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CCTGCTTTCCACATTCGAGGAGATGCTTGTTCGAAAGAAAATGAAATGTTGATGTGAGT 1545
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTACGATATATAACAATGAGTCTCCATCAACCATCTTCTGAAGCTCAAGAGAAATCC 1605
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 1606 AAAAGCCTTAAATTAATCTCAATATATGCAGGAGATGCTCTAAGAGAAATATACATTTGTT 1665
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1666 TCAGAACATGCACAAAGAGACCAACGTGAACACAGTGTCAAAATGAAGGAAGCTGAACAC 1725
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db 1726 ATGTATCAAAACGAAACAAAGATAATGTGAACAAAACACACTGAACAGCAGGAGTCTCTAG 1785
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
Db 1786 CAGAAATTAATTTCAACTACAAGCAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1845
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1846 CATAGAAGAGCTGACAAACAAAGCAAGATATGATATTCATTTCTTTGAGAGGAAA 1905
Qy 281 MetGlnHisIleAsnLeuLysGluLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAACCGTATATATCAATATGAAAAGAGAAAGAGCAAGCAAGCAAGAAACTCA 2016
RESULT 14
AR454005 2307 bp DNA linear PAT 20-FEB-2004
LOCUS Sequence 468 from patent US 6680197.
ACCESSION AR454005
VERSION AR454005.1 GI:42686795
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2307)
Hepier,W.T. and Henderson,R.A.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: US 6680197-A 468 20-JAN-2004;
FEATURES Location/Qualifiers
source 1..2307
/mol_type="genomic DNA"
```

```
Alignment Scores:
Pred. No.: 1-38e-115 Length: 2307
Score: 1632.00 Matches: 317
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-489-079-25 (1-317) x AR454005 (1-2307)
Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
Db 1066 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACACTCATGAAAATGAAAATTAATCTC 1125
Qy 21 LeuHisGluAsnCysMetLeuLysLysGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
Db 1126 TTACATGAAAATTTGCATGTTGAAAAGGAAATTCCTATGCTTAAACCTGSAATATAGCCACA 1185
Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
Db 1186 CTGAACACCCATACACAGGAAAGGAAATTAATATCTTTGAGGACATTAAGATTTTAAATA 1245
Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
Db 1246 GAAAAGAAATGCTGAACCTTCAGATGACCTTAAACCTGAAGAGGAATCATTTAACTAAAAAGG 1305
Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAenThrMetLeuThrSer 100
Db 1306 GCATCTCAATATATAGTGGGACGCTTAAAGTTTCTGATAGCTGAGAACACAATGCTCAGTTCT 1365
Qy 101 LysLeuLysGluLysGlnAspLysGluLeuLeuGluAlaGluLeuGluSerHisHisPro 120
Db 1366 AAATTGAAGGAAACAAAGCAAAAGAAATATCTAGAGGCAAGAAATTTGAATCACACCATCTCT 1425
Qy 121 ArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGlu 140
Db 1426 AGACTGGCTTCTGCTGTACAGACCATGATCAAAATTTGTGCATCAAGAAAAGTCAAGAA 1485
Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160
Db 1486 CCTGCTTTCCACATTCGAGGAGATGCTTGTTCGAAAGAAAATGAAATGTTGATGTGAGT 1545
Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180
Db 1546 AGTACGATATATAACAATGAGTGTCTCCATCAACCATCTTCTGAAGCTCAAGAGGAATCC 1605
Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200
Db 1606 AAAAGCCTTAAATTAATCTCAATTAATGCAGGAGATGCTCTAAGAGAAAATATACATTTGTT 1665
Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220
Db 1666 TCAGAACATGCACAAAGAGACCAACCGTGAACACAGTGTCAATGAAAGGAGCTGAACAC 1725
Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240
Db 1726 ATGTATCAAAACGAAACAAAGATAATGTGAACAAAACACACTGAACAGCAGGAGTCTCTAG 1785
Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGlnGlnLeuValHisAla 260
Db 1786 CAGAAATTAATTTCAACTACAAGCAAAATATGTGGCTTCAACAGCAATTAGTTTCATGCA 1845
Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280
Db 1846 CATAGAAGAGCTGACAAACAAAGCAAGATATGATATTCATTTCTTTGAGAGGAAA 1905
Qy 281 MetGlnHisIleAsnLeuLysGluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeu 300
Db 1906 ATGCACATCATCTCTTAAAGAGAAAATGAGGAGATATTTAATTACATAATCAACATTTA 1965
Qy 301 LysAsnArgIleTyrGlnTyrGluLysGluLysAlaGluThrGluAsnSer 317
Db 1966 AAAACCGTATATATCAATATGAAAAGAGAAAGAGCAAGCAAGCAAGAAACTCA 2016
```

Db 1966 AAAAACCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAAACCTCA 2016

RESULT 15

AR561593

LOCUS AR561593 2307 bp DNA linear PAT 08-OCT-2004

DEFINITION Sequence 468 from patent US 6756477.

ACCESSION AR561593

VERSION AR561593.1 GI:53974501

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1. (bases 1 to 2307)

AUTHORS Jiang, Y., Dillon, D.C., Mitcham, J.L., Xu, J. and Harlocker, S.L.

TITLE Compositions and methods for the therapy and diagnosis of breast cancer

JOURNAL Patent: US 6756477-A 468 29-JUN-2004;

FEATURES

source Location/Qualifiers

1..2307

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 1.38e-115 Length: 2307

Score: 1632.00 Matches: 317

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-489-079-25 (1-317) x AR561593 (1-2307)

Qy 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20

Db 1066 ATGGGAACAAGAGCTCTGCAGTGTGAGGTTTCTCACATCATGAAATGAAATTTATCTC 1125

Qy 21 LeuHisGluAsnCysMetLeuLysLysGluIleAlaMetLeuLysLeuGluIleAlaThr 40

Db 1126 TTACATGAAATTTGCATGTTGAAAGGAAATTCATGCTTAAACTGAAATAGCCACA 1185

Qy 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60

Db 1186 CTGAAACACCAATACACAGGAAAGGAAATAATAATCTTTGAGGACATTAAGATTTTAA 1245

Qy 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80

Db 1246 GAAAGAATGCTGAACCTTCAGATGACCTTAAACCTGAAAGAGAAATCATTAACCTAAAGG 1305

Qy 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuIleAlaGluAsnThrMetLeuThrSer 100

Db 1306 GCATCTCAATATAGTGGCAGCTTAAAGTTCTGATAGCTGAGACACACAATGCTCACTTCT 1365

Qy 101 LysLeuLysGluLysGlnAspLysGluIleLeuGluAlaGluIleGluSerHisPro 120

Db 1366 AAATTGAAGGAAAAACAACAGAAATACTAGAGGCGAGAAATGAAATTCACACCATCCT 1425

Qy 121 ArgLeuAlaSerAlaValGlnAspHisGlnIleValThrSerArgLysSerGlnGlu 140

Db 1426 AGACTGGCTTCGTGTACAGACCATGATCAATTTGTGACATCAAGAAAAAGTCAAGAA 1485

Qy 141 ProAlaPheHisIleAlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSer 160

Db 1486 CTGCTTTCCACATGTCAGAGATGCTTGTGCAAGAAAAAATGAAATGTTGATGTGAGT 1545

Qy 161 SerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSer 180

Db 1546 AGTACGATATAAACAATAGGTGCTCCATCAACCACTTTCTGAAAGCTCAAAAGGAAATCC 1605

Qy 181 LysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGluAsnThrLeuVal 200

Db 1606 AAAAGCCTAAAAATTAATCTCAATTATGAGGAGATGCTCTAAGAGAAAAATACATTGGTT 1665

Qy 201 SerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysGluAlaGluHis 220

Db 1666 TCAGAACATGCACAAAGAGACCAACGCGAAACACACAGTGTCAATGAAAGAGCTGAAACAC 1725

Qy 221 MetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGluGlnGlnGluSerLeuAsp 240

Db 1726 ATGTATCAAAACGAACAAGATAATGTGAACAAACACACTGAACACAGCAGGAGTCTCTAGAT 1785

Qy 241 GlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnGlnGlnLeuValHisAla 260

Db 1786 CAGAAATTTATTTCAACTACAAAGCAAAAATATGTGGCTTCAACACGCAATTAGTTTCATGCA 1845

Qy 261 HisLysLysAlaAspAsnLysSerLysIleThrIleAspIleHisPheLeuGluArgLys 280

Db 1846 CATAGAAAGCTGACAAACAAAGCAAGATAACAATTGATATTCATTTTCTTGAGAGGAAA 1905

Qy 281 MetGlnHisHisLeuLeuLysGluLysGluLysGluLysGluLysPheAsnTyrAsnAsnHisLeu 300

Db 1906 ATGCAACATCATCTCTTAAAGAGAAAATGAGGAGATATTTAATTACAATAACCATTTA 1965

Qy 301 LysAsnArgIleTyrGlnTyrGlnLysGluLysGluLysAlaGluThrGluAsnSer 317

Db 1966 AAAAACCGTATATATCAATATGAAAGAGAGAAAGCAGAAACAGAAAACCTCA 2016

Search completed: August 2, 2005, 02:51:17

Job time : 4815.32 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:04:01 ; Search time 112.678 Seconds
(without alignments)
1088.086 Million cell updates/sec

Title: US-09-489-079-25
Perfect score: 1632
Sequence: 1 MGTRALQCEVSHTHENYL.....NHLKNRIYQYKEKAETENS 317

Scoring table: BLOSUM62 / Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 16Dec04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1632	100.0	317	3 AAB07639	Aab07639 Amino aci
2	1622	99.4	317	4 AAB50261	Aab50261 Human bre
3	1622	99.4	317	4 AAG65980	Aag65980 B726P fir
4	1622	99.4	317	4 AAU33343	Aau33343 Human bre
5	1622	99.4	317	5 AAOL5777	Aaol5777 Human bre
6	1622	99.4	317	5 ABG78910	Abg78910 Human bre
7	1622	99.4	317	6 ABJ37733	Abj37733 Human tum
8	1622	99.4	317	7 ADL92838	Adl92838 Human bre
9	1622	99.4	317	8 ADE44128	Ade44128 Human bre
10	1622	99.4	323	4 AAU33353	Aau33353 Human pro
11	1622	99.4	323	5 ABG78920	Abg78920 Human bre
12	1622	99.4	323	6 ABJ37743	Abj37743 Human tum
13	1622	99.4	323	7 ADL93145	Adl93145 Human bre
14	1622	99.4	410	4 AAU33359	Aau33359 Human bre
15	1622	99.4	410	5 ABG78926	Abg78926 Human bre
16	1622	99.4	410	6 ABJ37749	Abj37749 Human tum
17	1622	99.4	410	7 ADL93157	Adl93157 Human bre
18	1617	99.1	324	7 ADL93213	Adl93213 Human tum
19	1617	99.1	324	7 AAB50244	Aab50244 Human bre
20	1586	97.2	432	4 AAG65982	Aag65982 B726P thi
21	1586	97.2	432	4 AAU33345	Aau33345 Human bre
22	1586	97.2	432	4 AAOL5779	Aaol5779 Human bre
23	1586	97.2	432	5 ABG78912	Abg78912 Human bre
24	1586	97.2	432	5 ABJ37735	Abj37735 Human tum
25	1586	97.2	432	6 ABJ37735	Abj37735 Human tum

26	1586	97.2	432	7 ADL92843	Adl92843 Human bre
27	1586	97.2	432	8 ADE44133	Ade44133 Human bre
28	1586	97.2	1002	4 AAU33351	Aau33351 Human bre
29	1586	97.2	1002	5 ABG78918	Abg78918 Human bre
30	1586	97.2	1002	6 ABJ37741	Abj37741 Human tum
31	1586	97.2	1002	7 ADL93137	Adl93137 Human bre
32	1586	97.2	1002	8 ADE44427	Ade44427 Human bre
33	1586	97.2	1095	4 AAU33357	Aau33357 Human bre
34	1586	97.2	1095	5 ABG78924	Abg78924 Human bre
35	1586	97.2	1095	6 ABJ37747	Abj37747 Human tum
36	1586	97.2	1095	7 ADL93155	Adl93155 Human bre
37	1586	97.2	1341	4 AAB84702	Aab84702 Amino aci
38	1586	97.2	1341	5 ABJ05537	Abj05537 Breast ca
39	1586	97.2	1341	6 ABJ37784	Abj37784 Human tum
40	1586	97.2	1341	6 ABR47548	AbR47548 Breast ca
41	1586	97.2	1341	7 ADL93227	Adl93227 Human bre
42	1576	96.6	1013	6 ABJ37783	Abj37783 Human tum
43	1576	96.6	1013	7 ADL93215	Adl93215 Human bre
44	1576	96.6	1349	6 ABJ37788	Abj37788 Human tum
45	1576	96.6	1349	7 ADL93235	Adl93235 Human bre

ALIGNMENTS

RESULT 1
AAB07639
ID AAB07639 standard; protein; 317 AA.

AC AAB07639;

XX 07-NOV-2000 (first entry)

XX Amino acid sequence of BS322 polypeptide.

XX BS322; breast tissue marker; breast disease; breast cancer.

XX Homo sapiens.

XX WO200043420-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-US001452.

XX 21-JAN-1999; 99US-00234716.

XX (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

XX Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;

XX Stroupe SD;

XX WPI; 2000-499217/44.

XX N-PSDB; AAAS9015.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the

XX detection and diagnosis of breast disease especially breast cancer.

XX Claim 23; Page 123-124; 126pp; English.

XX The present sequence represents a human BS322 polypeptide. BS322 is a

XX breast tissue marker. The BS322 polynucleotides and polypeptides are used

XX to detect and diagnose breast disease, e.g. breast cancer. The BS322

XX polynucleotides are useful as a source of probes and primers, and the

XX BS322 polypeptides are useful as antigens

XX Sequence 317 AA;

XX Query Match 100.0%; Score 1632; DB 3; Length 317;

XX Best Local Similarity 100.0%; Pred. No. 5.7e-118; Indels 0; Gaps 0;

XX Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60
Db 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60
Qy 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKLEABIESHHP 120
Db 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKLEABIESHHP 120
Qy 121 RLASAVODHDOI VTSRKSQEPAFHIAAGDACLQRKNVDVSTIYNNEVLHQP LSEAORKS 180
Db 121 RLASAVODHDOI VTSRKSQEPAFHIAAGDACLQRKNVDVSTIYNNEVLHQP LSEAORKS 180
Qy 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCCKEAEHMYQNEQDNVKNKHTEQQESLD 240
Db 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCCKEAEHMYQNEQDNVKNKHTEQQESLD 240
Qy 241 QKLFQLSQKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIPFNYNHL 300
Db 241 QKLFQLSQKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIPFNYNHL 300
Qy 301 KNRIYOYEKEKAETENS 317
Db 301 KNRIYOYEKEKAETENS 317

RESULT 2

AAB50261
ID AAB50261 standard; protein; 317 AA.
AC AAB50261;
DT 13-MAR-2001 (first entry)
DE Human breast cancer associated B726P-20 protein.
XX Human; breast cancer associated gene; vaccine; diagnosis; therapy.
KW Homo sapiens.
OS WO2000060076-A2.
PN 12-OCT-2000.
PD 15-FEB-2000; 2000WO-US005308.
PF 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
XX (CORI-) CORIXA CORP.
PA Yuqiu J, Dillon DC, Mitcham JL, Xu J, Harlocker SL;
PI WPI; 2001-122627/13.
DR N-PSDB; AAF17690.
XX An isolated polypeptide useful for the treatment and diagnosis of tumors
PT e.g. breast cancer comprises at least an immunogenic portion of a breast
PT tumor protein.
PS Claim 3; Page 142; 238pp; English.
XX The present invention provides the coding sequences and some protein
CC sequences of proteins associated with breast cancer in humans. These
CC sequences can be used in the diagnosis and treatment of cancers,
CC particularly breast tumours
XX
SQ Sequence 317 AA;

Query Match 99.4%; Score 1622; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60
Db 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60
Qy 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKLEABIESHHP 120
Db 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKLEABIESHHP 120
Qy 121 RLASAVODHDOI VTSRKSQEPAFHIAAGDACLQRKNVDVSTIYNNEVLHQP LSEAORKS 180
Db 121 RLASAVODHDOI VTSRKSQEPAFHIAAGDACLQRKNVDVSTIYNNEVLHQP LSEAORKS 180
Qy 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCCKEAEHMYQNEQDNVKNKHTEQQESLD 240
Db 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCCKEAEHMYQNEQDNVKNKHTEQQESLD 240
Qy 241 QKLFQLSQKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIPFNYNHL 300
Db 241 QKLFQLSQKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIPFNYNHL 300
Qy 301 KNRIYOYEKEKAETE 315
Db 301 KNRIYOYEKEKAETE 315
RESULT 3
AAG65980
ID AAG65980 standard; protein; 317 AA.
AC AAG65980;
DT 11-FEB-2002 (first entry)
DE B726P first splice variant.
XX Genetic subtraction; DNA microarray analysis; polymerase chain reaction;
KW cancer; B726P.
OS Homo sapiens.
PN WO200175171-A2.
PD 11-OCT-2001.
PF 02-APR-2001; 2001WO-US010631.
XX 03-APR-2000; 2000US-0194241P.
PR 20-JUL-2000; 2000US-0219862P.
PR 27-JUL-2000; 2000US-0221300P.
PR 18-DEC-2000; 2000US-0256592P.
XX (CORI-) CORIXA CORP.
PA Houghton RL, Dillon DC, Molesh DA, Xu J, Zehentner B, Persing DH;
PI WPI; 2001-626449/72.
DR N-PSDB; AAI67215.
XX Identifying tissue (tumor)-specific polynucleotides overexpressed in
PT tissue of interest as compared to control tissue, for detecting cancer
PT cells in patient, comprises DNA microarray analysis or quantitative
PT polymerase chain reaction.
XX Example; Page 101; 127pp; English.
PS The invention relates to identifying tissue-specific polynucleotides (P)
XX that involves performing a genetic subtraction to identify pool of (P)
CC from tissue of interest (Ti), performing DNA microarray analysis to
CC identify first subset of polynucleotides (SP1) at least 2-fold over
CC expressed in Ti, and performing quantitative polymerase chain reaction
CC (PCR) analysis on SP1 to identify second subset of (P). The method is
CC useful for determining the presence or absence of a cancer cell in a

CC patient, monitoring the progression of cancer in a patient using a
CC biological sample such as blood, serum, lymph nodes, bone marrow, sputum,
CC urine or a tumour biopsy sample. The methods are useful for determining
CC the presence or absence of or monitoring progression of prostate, breast,
CC colon, ovarian, lung, head and neck, lymphoma, leukemia, melanoma, liver,
CC gastric, kidney, bladder, pancreatic or endometrial cancer. The present
CC sequence represents the B726P first splice variant
XX
SQ Sequence 317 AA;

Query Match 99.4%; Score 1622; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117; Indels 0; Gaps 0;
Matches 315; Conservative 0; Mismatches 0;
QY 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILK 60
DB 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILK 60
QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHP 120
DB 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHP 120
QY 121 RLASAVQDHQIVTSRKQEPAFHAGDACLQKRNKNDVSVSTIYNNEVLHQPSEAQRKS 180
DB 121 RLASAVQDHQIVTSRKQEPAFHAGDACLQKRNKNDVSVSTIYNNEVLHQPSEAQRKS 180
QY 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKHTEQESILD 240
DB 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKHTEQESILD 240
QY 241 OKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKKEKNEEIFNYYNHL 300
DB 241 OKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKKEKNEEIFNYYNHL 300
QY 301 KNRYYQYKEKAETE 315
DB 301 KNRYYQYKEKAETE 315

RESULT 4
AAU33343
ID AAU33343 standard; protein; 317 AA.
AC AAU33343;
DT 18-DEC-2001 (first entry)
XX Human breast cancer protein encoded by cDNA cloneB726P-20.
DE Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy; immunogen.
XX Homo sapiens.
OS WO200179286-A2.
PN 25-OCT-2001.
PD 12-APR-2001; 2001WO-US012164.
PF 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX (CORI-) CORIXA CORP.
PA Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX WPI; 2001-611721/70.
XX N-PSDB; AAS47120.
XX Breast Tumor Proteins and nucleic acids useful for the prevention,

PT diagnosis and treatment of breast cancer.
XX Claim 3; Page 197; 297pp; English.
PS The invention relates to isolated breast tumour proteins and nucleic
XX acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
XX
SQ Sequence 317 AA;

Query Match 99.4%; Score 1622; DB 4; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117; Indels 0; Gaps 0;
Matches 315; Conservative 0; Mismatches 0;
QY 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILK 60
DB 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENYFEDIKILK 60
QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHP 120
DB 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLKEKQDKEILEAEIESHHP 120
QY 121 RLASAVQDHQIVTSRKQEPAFHAGDACLQKRNKNDVSVSTIYNNEVLHQPSEAQRKS 180
DB 121 RLASAVQDHQIVTSRKQEPAFHAGDACLQKRNKNDVSVSTIYNNEVLHQPSEAQRKS 180
QY 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKHTEQESILD 240
DB 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKHTEQESILD 240
QY 241 OKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKKEKNEEIFNYYNHL 300
DB 241 OKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKKEKNEEIFNYYNHL 300
QY 301 KNRYYQYKEKAETE 315
DB 301 KNRYYQYKEKAETE 315

RESULT 5
AAO15777
ID AAO15777 standard; protein; 317 AA.
XX AAO15777;
AC AAO15777;
DT 05-DEC-2002 (first entry)
XX Human breast antigen protein #2.
DE Immunogenic protein; breast antigen; breast antigen-related DNA; vaccine;
KW breast cancer; adoptive immunotherapy; tumour-reactive T-cell; cancer.
XX Homo sapiens.
OS US2002102602-A1.
PN

```
XX PD 01-AUG-2002.
XX PF 23-JUN-1999; 99US-00339338.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PA (YUJI/) YUJIU J.
XX PA (DILL/) DILLON D C.
XX PA (MITC/) MITCHAM J L.
XX PA (XUJJ/) XU J.
XX PI Yuqiu J, Dillon DC, Mitcham JL, Xu J;
XX DR WPI; 2002-697821/75.
XX DR N-PSDB; ABT08775.
XX PT Novel isolated polypeptide comprising an immunogenic portion of a breast
XX PT antigen, useful for treating breast cancer or for inhibiting the
XX PT development of breast cancer in a patient.
XX PS Claim 2; Page 77; 127pp; English.
XX CC The invention comprises a protein which contains an immunogenic portion
XX CC of a breast antigen. The invention also comprises breast antigen-related
XX CC DNA sequences. The protein of the invention is useful as a vaccine for
XX CC inhibiting the development of breast cancer. The protein of the invention
XX CC is used in adoptive immunotherapy for the treatment of cancer, and
XX CC generating/isolating tumour-reactive T-cells which can be administered to
XX CC a patient. The present amino acid sequence represents a human breast
XX CC antigen protein of the invention
XX SQ Sequence 317 AA;
Query Match 99.4%; Score 1622; DB 5; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGTALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHOYQKENKYPEDIKILK 60
Db 1 MGTALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHOYQKENKYPEDIKILK 60
Qy 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEKQDKETLEAEIESHHP 120
Db 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEKQDKETLEAEIESHHP 120
Qy 121 RLASAVQDHDQIVTSRKSQEPAFHAGDACLQRMNVDSVSTIYNNEVLHQPPLSEAQRKS 180
Db 121 RLASAVQDHDQIVTSRKSQEPAFHAGDACLQRMNVDSVSTIYNNEVLHQPPLSEAQRKS 180
Qy 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTEQQSLSLD 240
Db 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTEQQSLSLD 240
Qy 241 OKLFQLOSKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIFNYYNHL 300
Db 241 OKLFQLOSKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIFNYYNHL 300
Qy 301 KNRIYQYEKEKAETE 315
Db 301 KNRIYQYEKEKAETE 315
RESULT 6
ABG78910
ID ABG78910 standard; protein; 317 AA.
XX AC ABG78910;
XX XX 15-NOV-2002 (first entry)
XX DE Human breast tumour polypeptide #2.
```

```
XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX OS Homo sapiens.
XX PN US2002085998-A1.
XX PD 04-JUL-2002.
XX PF 13-APR-2001; 2001US-00834759.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PA (CORI-) CORIXA CORP.
XX XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX PI Henderson RA;
XX XX WPI; 2002-635657/68.
XX DR N-PSDB; ABS63721.
XX PT Novel breast cancer polynucleotides and polypeptides encoded by the
XX PT polynucleotides, useful for detecting the presence of breast cancer in a
XX PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX PS Claim 2; Page 106-107; 247pp; English.
XX CC The invention relates to an isolated breast tumour polynucleotide and the
XX CC polypeptide it encodes. The polynucleotide and polypeptide are useful for
XX CC detecting the presence of breast cancer in a patient, and in
XX CC pharmaceutical compositions for treating breast cancer. The sequences are
XX CC useful for stimulating an immune response in a patient and can therefore
XX CC be used in production of vaccines. The sequences are also useful for
XX CC detecting the presence of a cancer in a patient, by obtaining a
XX CC biological sample from the patient, contacting the biological sample with
XX CC a composition of the invention and detecting the amount of polynucleotide
XX CC that hybridizes to the sample. This sequence represents a human breast
XX CC tumour polypeptide of the invention
XX SQ Sequence 317 AA;
Query Match 99.4%; Score 1622; DB 5; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGTALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHOYQKENKYPEDIKILK 60
Db 1 MGTALQCEVSHTHENENYLLHNCMLKKEIAMLKLEIATLKHOYQKENKYPEDIKILK 60
Qy 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEKQDKETLEAEIESHHP 120
Db 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEKQDKETLEAEIESHHP 120
Qy 121 RLASAVQDHDQIVTSRKSQEPAFHAGDACLQRMNVDSVSTIYNNEVLHQPPLSEAQRKS 180
Db 121 RLASAVQDHDQIVTSRKSQEPAFHAGDACLQRMNVDSVSTIYNNEVLHQPPLSEAQRKS 180
Qy 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTEQQSLSLD 240
Db 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCMKEAEHMYQNEQDNVKNKHTEQQSLSLD 240
Qy 241 OKLFQLOSKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIFNYYNHL 300
Db 241 OKLFQLOSKNMWLQQQLVHAHKADNKSKITIDHFLERKMQHLLKKEKNEEIFNYYNHL 300
```


QY 301 KNRIYQYEKEKAETE 315
| | | | | | | | | |
Db 301 KNRIYQYEKEKAETE 315
| | | | | | | | | |
RESULT 7
ABJ37733
ID ABJ37733 standard; protein; 317 AA.
XX AC ABJ37733;
XX 15-MAY-2003 (first entry)
DE Human tumour-related protein - SEQ ID No 176.
KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
KW tumour; breast cancer; cancer; immune response stimulation.
XX OS Homo sapiens.
XX WO200283956-A1.
XX PD 24-OCT-2002.
XX PF 15-APR-2002; 2002WO-US012378.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
PI Mitchem JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Ranger GR;
PI Vedwick TS, McNeill PD, Durham M;
XX WPI; 2003-103376/09.
DR New polypeptide and polynucleotide useful for stimulating and/or
PT expanding T cells specific for a tumor protein and treating breast
PT cancer.
XX Example 1; Page 207-208; 375pp; English.
XX The invention comprises a method of stimulating and/or expanding T cells
CC specific for a tumour protein. The invention further comprises human
CC nucleic acids and proteins that are associated with tumours (e.g. breast
CC cancer). The method and sequences of the invention are useful for
CC stimulating and/or expanding T cells specific for a tumour protein,
CC detecting the presence of cancer, stimulating an immune response in a
CC patient and treating breast cancer. The present amino acid sequence
CC represents a human tumour-related protein
XX Sequence 317 AA;
Query Match 99.4%; Score 1622; DB 6; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117; Mismatches 0; Gaps 0;
Matches 315; Conservative 0; Indels 0; Gaps 0;
QY 1 MGTALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILK 60
| | | | | | | | | |
Db 1 MGTALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILK 60
| | | | | | | | | |
QY 61 EKNAELQNTLKLKEESLITRASQYSQLVLAENTMLTSKLKEDKDEILAEIESHP 120
| | | | | | | | | |
Db 61 EKNAELQNTLKLKEESLITRASQYSQLVLAENTMLTSKLKEDKDEILAEIESHP 120
| | | | | | | | | |
QY 121 RLASAVQDDHQIVTSRKSQBPAPHIAGDACLQKKNVDVSSITYNNEVLHQLPSEARQKS 180
| | | | | | | | | |
Db 121 RLASAVQDDHQIVTSRKSQBPAPHIAGDACLQKKNVDVSSITYNNEVLHQLPSEARQKS 180
| | | | | | | | | |
QY 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCMKEAHMYQNEQDNVNVKHTEQEESLD 240
| | | | | | | | | |

Db 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCMKEAHMYQNEQDNVNVKHTEQEESLD 240
| | | | | | | | | |
QY 241 QKLFQLOSNNWLOQQLVHAHKADNKSKITIDIHFLERKNQHHLLKKEKBEIFNYYNNHL 300
| | | | | | | | | |
Db 241 QKLFQLOSNNWLOQQLVHAHKADNKSKITIDIHFLERKNQHHLLKKEKBEIFNYYNNHL 300
| | | | | | | | | |
QY 301 KNRIYQYEKEKAETE 315
| | | | | | | | | |
Db 301 KNRIYQYEKEKAETE 315
| | | | | | | | | |
RESULT 8
ADL92838
ID ADL92838 standard; protein; 317 AA.
XX AC ADL92838;
XX 20-MAY-2004 (first entry)
DE Human breast cancer-associated polypeptide #2.
KW Human breast cancer-associated polypeptide #2.
KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX OS Homo sapiens.
XX US200316022-A1.
XX PD 04-SEP-2003.
XX PF 15-APR-2002; 2002US-00124805.
XX PR 28-DEC-1998; 98US-00222575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX (CORI-) CORIXA CORP.
XX Houghton RL, Sleath PR, Persing DH;
XX WPI; 2003-874918/81.
XX N-PSDB; ADL92837.
PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
PT treating breast cancer.
XX Example 1; SEQ ID NO 176; 294pp; English.
XX The invention relates to an isolated breast cancer-associated
CC polypeptide. The polypeptide may be used for the diagnosis and treatment
CC of breast cancers. The methods are useful for detecting the presence of a
CC cancer in a patient and treating a cancer in a patient. The present
CC sequence represents the amino acid sequence of a human breast cancer-
CC associated polypeptide.
XX Sequence 317 AA;
Query Match 99.4%; Score 1622; DB 7; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.4e-117; Mismatches 0; Gaps 0;
Matches 315; Conservative 0; Indels 0; Gaps 0;
QY 1 MGTALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILK 60
| | | | | | | | | |
Db 1 MGTALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKFEDIKILK 60
| | | | | | | | | |

XX PS Example 6; Page 286-287; 297pp; English.

XX CC The invention relates to isolated breast tumour proteins and nucleic

XX CC acids that encode them, including immunogenic fragments of the proteins.

XX CC Also included are expression vectors expressing the proteins, transformed

XX CC cells and antibodies raised against the proteins or an antigen presenting

XX CC cell expressing the protein. The proteins and nucleic acids may be used

XX CC in the prevention, diagnosis and treatment of diseases associated with

XX CC inappropriate breast tumour protein expression, i.e. breast tumours and

XX CC breast cancer e.g. by gene therapy. The nucleic acids and their

XX CC complements may also be used as DNA probes in diagnostic assays to detect

XX CC and quantitate the presence of similar nucleic acids in samples, and

XX CC therefore which patients may be in need of restorative therapy. The

XX CC proteins, nucleic acids and antibodies may be used in assays to identify

XX CC modulators (e.g. antagonists) of breast tumour protein expression and

XX CC activity. The antibodies and antagonists may also be used to down

XX CC regulate expression and activity. The antibodies may also be used as

XX CC diagnostic agents for detecting the presence of the proteins in samples

XX CC (e.g. by enzyme linked immunosorbant assay (ELISA)) and in other immuno-

XX CC purification diagnostic techniques. The present sequence is a breast

XX CC tumour protein encoded by a cDNA from a breast tumour cDNA library

XX CC isolated by subtractive hybridisation against a normal breast cDNA

XX CC library

XX PS Sequence 323 AA;

Query Match 99.4%; Score 1622; DB 4; Length 323;

Best Local Similarity 100.0%; Pred. No. 3.5e-117; Indels 0; Gaps 0;

Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60

DB 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60

QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKKEILEAETESHHP 120

DB 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKKEILEAETESHHP 120

QY 121 RLASAVQDHQIVTSRKSOEPAFHTAGDACLQKKNVDVSSITYNNEVLHQPLSAQRKS 180

DB 121 RLASAVQDHQIVTSRKSOEPAFHTAGDACLQKKNVDVSSITYNNEVLHQPLSAQRKS 180

QY 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKTEQOESLD 240

DB 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKTEQOESLD 240

QY 241 QKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKEKNESIFNYYNNHL 300

DB 241 QKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKEKNESIFNYYNNHL 300

QY 301 KNRIIYQEKKEAETE 315

DB 301 KNRIIYQEKKEAETE 315

RESULT 11

ABG78920

ID ABG78920 standard; protein; 323 AA.

XX AC ABG78920;

XX DT 15-NOV-2002 (first entry)

XX DE Human breast tumour polypeptide #12.

XX KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.

XX OS Homo sapiens.

XX PN US2002085998-A1.

XX OS 04-JUL-2002.

XX PF 13-APR-2001; 2001US-00834759.

XX PR 28-DEC-1999; 98US-00222575.

XX PR 02-APR-1999; 99US-00285480.

XX PR 23-JUN-1999; 99US-00339338.

XX PR 02-SEP-1999; 99US-00389881.

XX PR 03-NOV-1999; 99US-00433826.

XX PR 17-APR-2000; 2000US-00551621.

XX PR 08-JUN-2000; 2000US-00590751.

XX PR 22-JUN-2000; 2000US-00604287.

XX PR 20-JUL-2000; 2000US-00620405.

XX PA (CORI-) CORIXA CORP.

PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;

PI Henderson RA;

XX WPI; 2002-635657/68.

DR N-PSDB; ABS64018.

XX Novel breast cancer polynucleotides and polypeptides encoded by the

PT polynucleotides, useful for detecting the presence of breast cancer in a

PT patient, and in pharmaceutical compositions, for treating breast cancer.

XX Example 7; Page 215-216; 247pp; English.

XX The invention relates to an isolated breast tumour polynucleotide and the

CC polypeptide it encodes. The polynucleotide and polypeptide are useful for

CC detecting the presence of breast cancer in a patient and in

CC pharmaceutical compositions for treating breast cancer. The sequences are

CC useful for stimulating an immune response in a patient and can therefore

CC be used in production of vaccines. The sequences are also useful for

CC detecting the presence of a cancer in a patient, by obtaining a

CC biological sample from the patient, contacting the biological sample with

CC a composition of the invention and detecting the amount of polynucleotide

CC that hybridizes to the sample. This sequence represents a human breast

XX tumour polypeptide of the invention

XX PS Sequence 323 AA;

Query Match 99.4%; Score 1622; DB 5; Length 323;

Best Local Similarity 100.0%; Pred. No. 3.5e-117; Indels 0; Gaps 0;

Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60

DB 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIAMLKLEIATLKHQYQEKENKYFEDIKILK 60

QY 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKKEILEAETESHHP 120

DB 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSLKKEQDKKEILEAETESHHP 120

QY 121 RLASAVQDHQIVTSRKSOEPAFHTAGDACLQKKNVDVSSITYNNEVLHQPLSAQRKS 180

DB 121 RLASAVQDHQIVTSRKSOEPAFHTAGDACLQKKNVDVSSITYNNEVLHQPLSAQRKS 180

QY 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKTEQOESLD 240

DB 181 KSLKINLYAGDALRENTLVSEHAQRDQRETQCKMKEAHMYQNEQDNVKNKTEQOESLD 240

QY 241 QKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKEKNESIFNYYNNHL 300

DB 241 QKLFQSQKNMWLQQLVHAHKADNKSKITIDIHFLERKMQHLLKEKNESIFNYYNNHL 300

QY 301 KNRIIYQEKKEAETE 315

DB 301 KNRIIYQEKKEAETE 315

RESULT 12

ABJ37743

ID ABJ37743 standard; protein; 323 AA.

```
XX AC ABJ37743;
XX AC 15-MAY-2003 (first entry)
XX DT Human tumour-related protein - SEQ ID NO 483.
XX DE Human; vaccine; gene therapy; T cell stimulation; T cell expansion;
XX KW tumour; breast cancer; cancer; immune response stimulation.
XX KW Homo sapiens.
XX OS WO200283956-A1.
XX FN 24-OCT-2002.
XX PD 15-APR-2002; 2002WO-US012378.
XX PF 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX PR (CORI-) CORIXA CORP.
XX PA Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;
XX PI Mitcham JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;
XX PI Vedvick TS, McNeill PD, Durham M;
XX PR WPI; 2003-103376/09.
XX DR New polypeptide and polynucleotide useful for stimulating and/or
XX PT expanding T cells specific for a tumor protein and treating breast
XX PT cancer.
XX PR Example 7; Page 303-304; 375pp; English.
XX PS The invention comprises a method of stimulating and/or expanding T cells
XX CC specific for a tumour protein. The invention further comprises human
XX CC nucleic acids and proteins that are associated with tumours (e.g. breast
XX CC cancer). The method and sequences of the invention are useful for
XX CC stimulating and/or expanding T cells specific for a tumour protein,
XX CC detecting the presence of cancer, stimulating an immune response in a
XX CC patient and treating breast cancer. The present amino acid sequence
XX CC represents a human tumour-related protein
XX SQ Sequence 323 AA;
XX Query Match 99.4%; Score 1622; DB 6; Length 323;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-117;
XX Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKLK 60
Db 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKLK 60
Qy 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEQKKEILEAIESHHP 120
Db 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEQKKEILEAIESHHP 120
Qy 121 RLASAVQDHQIVTSRKSQEPAFHAGDACLQKRMNVDSSTIYNNEVLHQPULSEAQRKS 180
Db 121 RLASAVQDHQIVTSRKSQEPAFHAGDACLQKRMNVDSSTIYNNEVLHQPULSEAQRKS 180
Qy 181 KSLKINLYAGDALRENTLSEHAQRQRETOCQKKEAHEMYQNEQNVNKHTEQESLSD 240
Db 181 KSLKINLYAGDALRENTLSEHAQRQRETOCQKKEAHEMYQNEQNVNKHTEQESLSD 240
Qy 241 QKLFQLQSKNNWLOQQLVHAHKKADNKSKITIDHFLERKMQHLLKEKNEEIFYNNHLL 300
Db 241 QKLFQLQSKNNWLOQQLVHAHKKADNKSKITIDHFLERKMQHLLKEKNEEIFYNNHLL 300
Qy 301 KNRIYQYEKEKAETE 315
|||||
```

```
Db 301 KNRIYQYEKEKAETE 315
RESULT 13
ADL93145
ID ADL93145 standard; protein; 323 AA.
XX AC ADL93145;
XX DT 20-MAY-2004 (first entry)
XX DE Human breast cancer-associated polypeptide #12.
XX KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.
XX OS Homo sapiens.
XX FN US2003166022-A1.
XX PD 04-SEP-2003.
XX PF 15-APR-2002; 2002US-00124805.
XX PR 28-DEC-1998; 98US-0022575.
XX PR 02-APR-1999; 99US-00285480.
XX PR 23-JUN-1999; 99US-00339338.
XX PR 02-SEP-1999; 99US-00389681.
XX PR 03-NOV-1999; 99US-00433826.
XX PR 17-APR-2000; 2000US-00551621.
XX PR 08-JUN-2000; 2000US-00590751.
XX PR 22-JUN-2000; 2000US-00604287.
XX PR 20-JUL-2000; 2000US-00620405.
XX PR 13-APR-2001; 2001US-00834759.
XX PR 07-DEC-2001; 2001US-00007805.
XX PR 13-FEB-2002; 2002US-00076622.
XX PA (CORI-) CORIXA CORP.
XX PI Houghton RL, Sleath PR, Persing DH;
XX DR WPI; 2003-874918/81.
XX DR N-PSDB; ADL93144.
XX PT An isolated oncogenic polypeptide useful for preventing, diagnosing and
XX PT treating breast cancer.
XX PS Example 7; SEQ ID NO 483; 294pp; English.
XX CC The invention relates to an isolated breast cancer-associated
XX CC polypeptide. The polypeptide may be used for the diagnosis and treatment
XX CC of breast cancers. The methods are useful for detecting the presence of a
XX CC cancer in a patient and treating a cancer in a patient. The present
XX CC sequence represents the amino acid sequence of a human breast cancer-
XX CC associated polypeptide.
XX SQ Sequence 323 AA;
XX Query Match 99.4%; Score 1622; DB 7; Length 323;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-117;
XX Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKLK 60
Db 1 MGTRALQCEVSHTHENENYLLHENCMLKKEIATLKHQYQEKENKYPEDIKLK 60
Qy 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEQKKEILEAIESHHP 120
Db 61 EKNAELQMTLKLKEESLTTRASQYSGQLKVLIAENTMLTSKLKEQKKEILEAIESHHP 120
Qy 121 RLASAVQDHQIVTSRKSQEPAFHAGDACLQKRMNVDSSTIYNNEVLHQPULSEAQRKS 180
Db 121 RLASAVQDHQIVTSRKSQEPAFHAGDACLQKRMNVDSSTIYNNEVLHQPULSEAQRKS 180
```

QY 181 KSLKINLYAGDALRENTLVSEHAQDQRETCQCKEAEHMYQNEQDNNVKNKHTEQCESLD 240
Db 181 KSLKINLYAGDALRENTLVSEHAQDQRETCQCKEAEHMYQNEQDNNVKNKHTEQCESLD 240
QY 241 QKLFOLOSKNNWLOOQLVHAHKADNKSKITIDHFLERKQHHLLKEKNEEIFNYYNHL 300
Db 241 QKLFOLOSKNNWLOOQLVHAHKADNKSKITIDHFLERKQHHLLKEKNEEIFNYYNHL 300
QY 301 KNRIYQYEKEKAETE 315
Db 301 KNRIYQYEKEKAETE 315

RESULT 14
AAU33359
ID AAU33359 standard; protein; 410 AA.
XX AAU33359;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human breast cancer protein B726P fusion protein #3.
XX
KW Human; ss; breast cancer protein; tumour; cancer; cytostatic;
KW gene therapy; immunogen.
XX
OS Homo sapiens.
XX
PN WO200179286-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US012164.
XX
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
DR WPI; 2001-611721/70.
DR N-PSDB; AAS47423.
XX
PT Breast Tumor Proteins and nucleic acids useful for the prevention,
PT diagnosis and treatment of breast cancer.
XX
PS Claim 24; Page 296-297; 297pp; English.
XX
CC The invention relates to isolated breast tumour proteins and nucleic
CC acids that encode them, including immunogenic fragments of the proteins.
CC Also included are expression vectors expressing the proteins, transformed
CC cells and antibodies raised against the proteins or an antigen presenting
CC cell expressing the protein. The proteins and nucleic acids may be used
CC in the prevention, diagnosis and treatment of diseases associated with
CC inappropriate breast tumour protein expression, i.e. breast tumours and
CC breast cancer e.g. by gene therapy. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to detect
CC and quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC proteins, nucleic acids and antibodies may be used in assays to identify
CC modulators (e.g. antagonists) of breast tumour protein expression and
CC activity. The antibodies and antagonists may also be used to down
CC regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the proteins in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)) and in other immuno-
CC purification diagnostic techniques. The present sequence is a breast
CC tumour protein encoded by a cDNA from a breast tumour cDNA library
CC isolated by subtractive hybridisation against a normal breast cDNA
CC library
XX

SQ Sequence 410 AA;
Query Match 99.4%; Score 1622; DB 4; Length 410;
Best Local Similarity 100.0%; Pred. No. 4.7e-117;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGTALQCEVSHTHENENYLLHENCMLKKEIATLKHQYQENKVFEDIKILK 60
Db 94 MGTALQCEVSHTHENENYLLHENCMLKKEIATLKHQYQENKVFEDIKILK 153
QY 61 EKNAELQMTLKEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKKEILEAETESHHP 120
Db 154 EKNAELQMTLKEESLTKRASQYSGQLKVLIAENTMTLSKLEKQDKKEILEAETESHHP 213
QY 121 FLASAVQDHDQIVTSRKSQEPAFHAGDACIQRKQNVDSVSTIYNNVHLPSEAQRKS 180
Db 214 FLASAVQDHDQIVTSRKSQEPAFHAGDACIQRKQNVDSVSTIYNNVHLPSEAQRKS 273
QY 181 KSLKINLYAGDALRENTLVSEHAQDQRETCQCKEAEHMYQNEQDNNVKNKHTEQCESLD 240
Db 274 KSLKINLYAGDALRENTLVSEHAQDQRETCQCKEAEHMYQNEQDNNVKNKHTEQCESLD 333
QY 241 QKLFOLOSKNNWLOOQLVHAHKADNKSKITIDHFLERKQHHLLKEKNEEIFNYYNHL 300
Db 334 QKLFOLOSKNNWLOOQLVHAHKADNKSKITIDHFLERKQHHLLKEKNEEIFNYYNHL 393
QY 301 KNRIYQYEKEKAETE 315
Db 394 KNRIYQYEKEKAETE 408

RESULT 15
ABG78926
ID ABG78926 standard; protein; 410 AA.
XX ABG78926;
AC ABG78926;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human breast tumour polypeptide #17.
XX
KW Human; breast tumour protein; breast cancer; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN US2002085998-A1.
XX
PD 04-JUL-2002.
XX
PF 13-APR-2001; 2001US-00834759.
XX
PR 28-DEC-1998; 98US-00222575.
PR 02-APR-1999; 99US-00285480.
PR 23-JUN-1999; 99US-00339338.
PR 02-SEP-1999; 99US-00389681.
PR 03-NOV-1999; 99US-00433826.
PR 17-APR-2000; 2000US-00551621.
PR 08-JUN-2000; 2000US-00590751.
PR 22-JUN-2000; 2000US-00604287.
PR 20-JUL-2000; 2000US-00620405.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
PI Henderson RA;
XX
XX WPI; 2002-635657/68.
DR N-PSDB; ABS64024.
XX
PT Novel breast cancer polynucleotides and polypeptides encoded by the
PT polynucleotides, useful for detecting the presence of breast cancer in a
PT patient, and in pharmaceutical compositions, for treating breast cancer.
XX

PS Disclosure; Page 228-229; 247pp; English.

XX The invention relates to an isolated breast tumour polynucleotide and the

CC polypeptide it encodes. The polynucleotide and polypeptide are useful for

CC detecting the presence of breast cancer in a patient, and in

CC pharmaceutical compositions for treating breast cancer. The sequences are

CC useful for stimulating an immune response in a patient and can therefore

CC be used in production of vaccines. The sequences are also useful for

CC detecting the presence of a cancer in a patient, by obtaining a

CC biological sample from the patient, contacting the biological sample with

CC a composition of the invention and detecting the amount of polynucleotide

CC that hybridizes to the sample. This sequence represents a human breast

CC tumour polypeptide of the invention

XX

SQ Sequence 410 AA;

Query Match 99.4%; Score 1622; DB 5; Length 410;

Best Local Similarity 100.0%; Pred. No. 4.7e-117;

Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGTALQCEVSHTHENENYLLHENCMKKEIATLKHQYQEKENKYPEDIKILK 60

Db 94 MGTALQCEVSHTHENENYLLHENCMKKEIATLKHQYQEKENKYPEDIKILK 153

Qy 61 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKLEAEIESHHP 120

Db 154 EKNAELQMTLKLKEESLTKRASQYSGQLKVLIAENTMLTSKLEKQDKLEAEIESHHP 213

Qy 121 RLASAVQDHQIVTSRKSQEPAFHAGDACLQRKQNVDSSTIYNNEVLHQPULSEAQRKS 180

Db 214 RLASAVQDHQIVTSRKSQEPAFHAGDACLQRKQNVDSSTIYNNEVLHQPULSEAQRKS 273

Qy 181 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCOMKEAHHMYQNEQDNVKNKHTEQQESLD 240

Db 274 KSLKINLNYAGDALRENTLVSEHAQRDQRETQCOMKEAHHMYQNEQDNVKNKHTEQQESLD 333

Qy 241 QKLFQLOSKNNWLQOQLVHAHKKADNKSKITIDHFLEKQKHLLKEKNEEIFNYYNNHL 300

Db 334 QKLFQLOSKNNWLQOQLVHAHKKADNKSKITIDHFLEKQKHLLKEKNEEIFNYYNNHL 393

Qy 301 KNRIYQYEKEKAETE 315

Db 394 KNRIYQYEKEKAETE 408

Search completed: August 1, 2005, 22:18:04

Job time : 113.678 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:44:18 ; Search time 12056.9 Seconds
(without alignments)
1000.786 Million cell updates/sec

Title: US-09-489-079-25

Perfect score: 1632
Sequence: 1 MGRALQCEVSHTHENENYL.....NHLKNRIYQYKEKAKTENS 317

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+p2n.model -DEV=xlp
-Q=/cgm2_1/USPTO_spool_p/US09489079/runat_01082005_141250_13309/app_query.fasta_1.1635
-DB=EST -QFMT=tastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1.1.10973 @runat_01082005_141250_13309 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hcc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gssi:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	926	56.7	565	2 AW373574	AW373574 QV4-BT053
2	846	51.8	492	5 BX282243	BX282243 BX282243
3	647.5	39.7	650	2 BE177744	BE177744 RC1-HT059
4	647	39.6	4760	3 CR749347	CR749347 Homo sapi
5	565	34.6	3527	3 HSN601761	AL136793 Homo sapi
6	527	32.3	3794	3 BC036210	BC036210 Homo sapi
7	513.5	31.5	638	5 BU678550	BU678550 UI-CF-DUI
8	509.5	31.2	700	6 CB527418	CB527418 UI-M-FYO-
9	487.5	29.9	699	7 CN458404	CN458404 UI-M-HNO-

10	481.5	29.5	862	4	BC431116	BC431116 602498742
11	470	28.8	600	5	BP238820	BP238820 BP238820
c 12	420.5	25.8	5193	3	CR627244	CR627244 Homo sapi
c 13	417.5	25.6	1014	5	BQ926802	BQ926802 AGENCOURT
14	416	25.5	670	7	CN415409	CN415409 170005321
c 15	407.5	25.0	886	5	BQ437838	BQ437838 AGENCOURT
16	382	23.4	583	7	CF894061	CF894061 A0110H05-
17	377	23.1	633	5	BQ557842	BQ557842 H4048F01-
18	372	22.8	505	6	CR220242	CR220242 LAB019C03
19	368	22.5	857	7	CR788460	CR788460 DKF2p459B
20	367.5	22.5	751	6	CD000265	CD000265 AGENCOURT
21	365.5	22.4	678	5	BQ787286	BQ787286 1152907.Y
22	356.5	21.8	630	6	CB298279	CB298279 220016.re
c 23	344	21.1	568	2	BF516480	BF516480 UI-H-BW1-
24	343	21.0	555	9	CF750511	CF750511 t19f-ges-
25	338.5	20.7	856	6	CA793848	CA793848 AGENCOURT
26	334.5	20.5	597	6	CB228494	CB228494 AGENCOURT
27	331.5	20.3	574	1	AV607651	AV607651 AV607651
28	331	20.3	858	6	CD244895	CD244895 AGENCOURT
29	328.5	20.1	574	6	CB298280	CB298280 220016.re
30	327	20.0	652	8	BZ924428	BZ924428 CH240_117
31	324.5	19.9	753	5	BX711332	BX711332 BX711332
32	324	19.9	555	7	CN415410	CN415410 170006001
33	323	19.8	641	6	CB044222	CB044222 NISC 9C03
c 34	320	19.6	694	1	AI687645	AI687645 tp8207.X
35	308	18.9	1093	3	BC022380	BC022380 Homo sapi
36	301	18.4	659	7	CO737543	CO737543 SILT02C17
37	299	18.3	417	8	CC248789	CC248789 XK525 Bay
38	297	18.2	449	6	CA576255	CA576255 K0642A10-
39	297	18.2	494	6	CA576220	CA576220 K0641F05-
40	297	18.2	501	6	CA573691	CA573691 K0604D04-
41	297	18.2	554	7	CF911609	CF911609 A0618B08-
42	297	18.2	562	7	CF911209	CF911209 A0626B11-
43	297	18.2	591	7	CN660762	CN660762 A0709B10-
44	297	18.2	610	7	CF911426	CF911426 A0615B07-
45	297	18.2	630	6	CA574480	CA574480 K0616H07-

ALIGNMENTS

AW373574 QV4-BT0534-281299-053-a01 BT0534 Homo sapiens cDNA, mRNA sequence. EST 04-FEB-2000

DEFINITION QV4-BT0534-281299-053-a01 BT0534 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW373574

VERSION AW373574.1 GI:6878228

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 565)

HCGP <http://www.ludwig.org.br/ORESTES>.

AUTHORS The FAPESP/LICR Human Cancer Genome Project

TITLE Unpublished (1999)

JOURNAL

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-BT0534-281299-053-a01&t3=1999-12-28&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 565.

Location/Qualifiers

1..565

/organism="Homo sapiens"

/mol_type="mRNA"

FEATURES

source

```

/db xref="taxon:9606"
/dev stage="Adult"
/clone_lib="BT0534"
/notes="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1.24e-88 Length: 565
Score: 926.00 Matches: 184
Percent Similarity: 98.40% Conservative: 1
Best Local Similarity: 97.87% Mismatches: 3
Query Match: 56.74% Indels: 0
DB: 2 Gaps: 0

```

US-09-489-079-25 (1-317) x AW373574 (1-565)

```

Qy 6 LeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeuLeuHisGluAsnCys 25
Db 1 CTCGAATCGACGGTTTCTCACACTCATGAAATGAAAAATTATCTTACATGAAATGTC 60
Qy 26 MetLeuLysGluLeuLeuAlaMetLeuLysGluLeuLeuAlaThrLeuLysHisGlnTyr 45
Db 61 ATGTGTGAAAAGGAAATTTGGCATCTTAAACTGGAATAGCCACACGAAACACCAATAC 120
Qy 46 GlnGluLysGluAsnLysTyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGlu 65
Db 121 CAGGAAAGGAAATAATAACTTTGAGACACTTAAGATTTTAAAGAAAGAAATGCTGAA 180
Qy 66 LeuGlnMetThrLeuLysGluLeuGluSerLeuThrLysArgAlaSerGlnTyrSer 85
Db 181 CTTTCAGATGACCTTAAACTGAAAGAGGAATCATTAACAAAGGGCATCTCAATATAGT 240
Qy 86 GlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSerLysLeuLysGluLys 105
Db 241 GGGCAGCTTAAAGTTCTGTATAGCTGAGAACCAATGCTCACTTCTTAATTAAGGAAAAA 300
Qy 106 GlnAspLysGluLeuLeuGluAlaGluLeuGluSerHisHisProArgLeuAlaSerAla 125
Db 301 CAGACAAAGAAATACTAGAGCGAGAAATGAATCACACCATCTAGACTGGCTTCTGCT 360
Qy 126 ValGlnAspHisAspGlnIleValThrSerArgLysSerGlnGluProAlaPheHisIle 145
Db 361 GTACAAGACCATGATCAAAATGTGACATCAAGAAAAAGTCAAGAACCTGCTTTCCACATT 420
Qy 146 AlaGlyAspAlaCysLeuGlnArgLysMetAsnValAspValSerSerThrIleTyrAsn 165
Db 421 GCAGGATGCTGCTTTTGCAAGAAAAAATGAATGTTGATGTAGTAGTACGATATATAAC 480
Qy 166 AsnGluValLeuHisGlnProLeuSerGluAlaGlnArgLysSerLysSerLeuLysIle 185
Db 481 AATGAGGTGCTCCATCAACCACTTCTTGAGCTCAAGAGAAATCCAAAGACCTTAAAAATT 540
Qy 186 AsnLeuAsnTyrAlaGlyAspAla 193
Db 541 ACTCTCAATTATGCAGGAGATGCT 564

```

```

RESULT 2
BX282243
LOCUS
DEFINITION BX282243 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
EST 04-MAR-2003
IMAGP9981196637 ; IMAGE:2659290, mRNA sequence.
ACCESSION BX282243
VERSION BX282243.1 GI:28615040
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 492)
 Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
 Radelof,U., Schneider,D. and Korn,B.
 Human Unigeneset - RZPD3
 Unpublished (2003)
 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGP9981196637.
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Human Unigeneset - RZPD3 (RZPDLIB No.972)
 http://www.rzpd.de/CloneCards/cgi-
 bin/showLib.pl.cgi/responserlibNo=972 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de
 This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.

FEATURES

source

```

Location/Qualifiers
1..492
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGP9981196637 ; IMAGE:2659290"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI_CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
728408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaudo."

```

ORIGIN

```

Alignment Scores:
Pred. No.: 3.64e-80 Length: 492
Score: 846.00 Matches: 162
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 51.84% Indels: 0
DB: 5 Gaps: 0
US-09-489-079-25 (1-317) x BX282243 (1-492)
Qy 156 AsnValAspValSerSerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGlu 175
Db 3 AATGTTGATGTGAGTAGTACGATATATAACAATGAGGTGCTCCATCAACCACTTTCTGAA 62
Qy 176 AlaGlnArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArg 195
Db 63 GCTCAAGAGAAATCCAAAGCCTTAAATTAATCTCAATTTATGCGAGAGATGCTCTAAGA 122
Qy 196 GluAsnThrLeuValSerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMet 215
Db 123 GAAATACATGTTGGTTTTCAGAACATGCAACAAAGAGACCAACGTGAAACACAGTGTCAAATG 182
Qy 216 LysGluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGlnGln 235
Db 183 AAGGAACCTGAACACATGTATCAAAAGCAACAGATAATGTGAACAAACACACTGAACAG 242
Qy 236 GlnGluSerLeuAspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTyrLeuGlnGln 255
Db 243 CAGGAGTCTCTAGATCAGAAATTTATTTCAACTACAAAGCAAAATATATGTGGCTTTCAACAG 302

```



```

gene
1. .3527
/genes="DKFZp434A171"
CDS
114..2585
/genes="DKFZp434A171"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66727.1"
/db_xref="GI:12053099"
/db_xref="UniProt/TREMBL:Q9H0H6"
/tranlations="MKLFGPSRRGTOAGSIDIHVVTGSGYRDRSELOKIHRAAVKG
DAAEVRCCLARRGGDLDALDKQHRTHALACASGHVQVVTLLVNRXCQIDVCDKENET
PLIQVHCQEACAVILLEGANPNFADLYGNTHALYAVTSESTSLAEKLSHGAIHE
ALDKNNTPLLFAIICKEKMWFLKKCAASHAVDRRLRSALMLAVYDPSGVNVL
LKNIDVFAOMCGRAEDYAI SHLTKIQQQLLEHKKIKLKEKSDVGSDESAYS1
FHELVDSPASDDKOLNVATKOCVPEKVEPLGSSHEKGNRI VNGQEGEPKAPHS
LKPSTVEDPANKVAGRNQVOTLRAEQALPVASBEORHERSEKKQPOVKEGNTN
KSKIQLSENI CDSSTSAAGRLTQQRKIGKTYPOQPPKLEKHEDRCRLKQNEEKT
NNMVLKKNRHELEKEKQVEAKOLEPTVQSLDKMSKTARNTPNDPFNHEMK
GLMDENCILKADITAILROEICTMKNNDLLEKENYKLDKIIVETNAALREKIKLNEEM
ITETAPRYQOELNDLAENETRLNAELKEKESKRLEADIESYQSRLAAAIKSHSESV
KTERNLKLALETVDVSVQVEMSSAISKVKAENEFTEQLSETQIKFNALKDKFCXTR
DSLRKSLALETVDNLSDVQOQTOEMKEMYNQAEAKVNNSTGKNVCEERI CHLORE
NWLVLQDLDDVQKEDHEKEIVTNIQRFIESEKDLVLBEKSKLWNECDHLKESLFQ
YEREKTVGVVSVIKEDKIFQTSRKTI"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 6,48e-49 Length: 3527
Score: 565.00 Matches: 123
Percent Similarity: 61.92% Conservative: 64
Best Local Similarity: 40.73% Mismatches: 113
Query Match: 34.62% Indels: 2
DB: 3 Gaps: 2

US-09-489-079-25 (1-317) x HSM801761 (1-3527)

Qy 12 HisThrHisGluAenGluAenTyrLeuLeuHisGluAenCysMetLeuLysLysGluLeu 31
Db 1617 CATAATCATGAAGAAATGAAAGGCTCATGATGATGAAATTCATTTTGAAGGCAGATATT 1676

Qy 32 AlaMetLeuLysLeuGluLeuAlaThrLeuLysHisGlnTyrGlnGlnLysGluAenLys 51
Db 1677 GCTATACTCAGACAGGAATATGTACAAATGAAATGCAACTTGGGAAAGAAATATAA 1736

Qy 52 TyrPheGluAspIleLysLeuLeuLysGluLysAlaGluLeuGlnMetThrLeuLys 71
Db 1737 TATCTTAAGGACATTAATAATTTGTTAAAGAAACAATGCTGCCCTTGAAGATATATAA 1796

Qy 72 LeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysValLeu 91
Db 1797 CTCAATCAGGAATGATAACAGAAACAGCATTCCGGTATCAACAGAGCTTAATGATCTC 1856

Qy 92 IleAlaGluAenThrMetLeuThrSerLys---LeuLysGluLysGlnAspLysGluLeu 110
Db 1857 AAGGCTCAGAATACAGGCTCAATGCGCACTGTTGAAGGAAAGAAAGCAAGAAAGA 1916

Qy 111 LeuGluAlaGluIleGluSerHisPheProArgLeuAlaSerAlaValGlnAspHisAsp 130
Db 1917 CTGGAAGCTGACATTAATCTTATCAGCTAGACTGCTGCTGCTGCTATAGCAACACAGT 1976

Qy 131 GlnIleValThrSerArgLysSerGlnGluProAlaPheHisIleAlaGlyAspAlaCys 150
Db 1977 GAAAGTGTGAAACAGAAAGAAACCTTAAACTTCTTTAGAGAGAACACAGAGATTTCT 2036

Qy 151 LeuGlnArgLysMetAsnValAspValSerSerThrIleTyrAsnAsnGluValLeuHis 170
Db 2037 GTACAAGTAGAATGAGTTCTTCGTTATTTCCAAAGTAAAGCTGAGAATGAGTTCTTACT 2096

Qy 171 GlnProLeuSerGluAlaGlnArgLysSerLysSerLeuLysValIleAsnLeuAsnTyrAla 190
Db 2097 GAAACAATTTCTGAAACACAATAATTAATTCATTCCTTAAAGATAGTAGTTCTGTGAACA 2156

Qy 191 GlyAspAlaLeuArgGluAenThrLeuValSerGluHisAlaGlnArgAspGlnArgGlu 210

```

```

Db 2157 AGAGATAGTCTCAGAAAAAAGTCATTGGCTTTAGAAACGTGTACAAAACGACCTAAGCCAA 2216
Qy 211 ThrGlnCysGlnMetLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsn 230
Db 2217 ACACAGCAGCAACACAGCAAGATGAAGAGATGTATCAAAATGCAGAGCTAAGTGAAT 2276
Qy 231 LysHisThrGluGlnGlnGlnSerLeuAspGlnLysLeuPheGlnLeuGlnSerLysAsn 250
Db 2277 AATTCCACTGGAAGTGAAGTGTGTAGAGAGAGATGTGTCCCTCCCAACGTGAAAT 2336
Qy 251 MetTyrLeuGlnGlnGlnLeuValHisAlaHisLysLysAlaAspAsnLysSerLysIle 270
Db 2337 GCGTGGCTGTGACAGCAACTAGATGACGTTTCATCAGAAAGAGAGATCATTAAGAGATGA 2396
Qy 271 Thr---IleAspIleHisPheLeuGluArgLysMetGlnHisLysLeuLysGluLys 289
Db 2397 ACTAATATCCAAAGAGCGCTTTATTGAGATGAAAGAGACCTCGTGTAGAGAGAAA 2456
Qy 290 AsnGluGluIlePheAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyrGluLys 309
Db 2457 AGTAAAGCTAATGATGAATGTGATCATTTTAAAGAAAGTCTCTTTTCAGTATGAGAGA 2516
Qy 310 GluLys 311
Db 2517 GAGAAA 2522

RESULT 6
BC036210
LOCUS BC036210 3794 bp mRNA linear HTC 20-SEP-2002
DEFINITION Homo sapiens, clone IMAGE:5265638, mRNA.
ACCESSION BC036210
VERSION BC036210.1 GI:23242519
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3794)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-@email.nih.gov
Tissue Procurement: Miklos Falkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-sngc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 63 Row: e Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: frame shifted.
FEATURES
source
1. .3794
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5265638"
/tissue_type="Testis"

```

```
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Alignment Scores:
Pred. No.:      8.3e-45      Length:      3794
Score:          527.00      Matches:      143
Percent Similarity: 55.71%      Conservative: 43
Best Local Similarity: 43.60%      Mismatches:  119
Query Match:      32.29%      Indels:      25
DB:              3          Gaps:      7

US-09-489-079-25 (1-317) x BC036210 (1-3794)

QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAsnTyrLeu 20
   |||
   |||
   |||
Db 1924 CTGGCTCGAAGGAAAGGAATTCCTCAGATTCTCTCAGCCATGACCAAGTACAAGATCTG 1983
   |||
   |||
   |||
QY 21 LeuHisGluAsnCysMetLeuLysGlyGluLeuAlaMetLeuLysLeuGluLeuAlaThr 40
   |||
   |||
   |||
Db 1984 TTGCATGAANAATCGCATGTTCAGAGACGAANITGCCCTCAGACTGGAATATACACA 2043
   |||
   |||
   |||
QY 41 LeuLysHisGlnTyrGlnGlnLysGluAsnLysTyrPheGluAspIleLysIleLeuLys 60
   |||
   |||
   |||
Db 2044 ATAGAAAACACAGACGAGAAATAAAGAACTATTTTGAGACATT----- 2091
   |||
   |||
   |||
QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLysArg 80
   |||
   |||
   |||
Db 2092 -----GGTCACCTCGAAAGACAAATAAAACCGGAATGGGAAATATTAAACAATG 2142
   |||
   |||
   |||
QY 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
   |||
   |||
   |||
Db 2143 ATATCGAGTATCTGCACAGCTTGGTGTCCTGACAGCTGAGAAATACCATGCTCCATCT 2202
   |||
   |||
   |||
QY 101 LysLeu---LysGluLysGlnAspLysGluLeuLeuGluAlaGluLeuSerHisHis 119
   |||
   |||
   |||
Db 2203 CAACCTGGAGAGGAAACAAAGAAAGAAAGCCTGGAGACAGAAGTTGAATCACACCTG 2262
   |||
   |||
   |||
QY 120 ProArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGln 139
   |||
   |||
   |||
Db 2263 TGCACACTGGCTGCTGTTAGACGTGATTGTGAACAAAGTGAGACATCCAAAAGACAGCGTG 2322
   |||
   |||
   |||
QY 140 GluProAlaPheHisIleAlaGlyAspAlaCys-----LeuGlnArgLysMetAsnVal 157
   |||
   |||
   |||
Db 2323 GAACCTGCTTTCAGGACCAAAAGATGAATGGTTTCATTATTCGGAGAAATGATTTT 2382
   |||
   |||
   |||
QY 158 AspValSerSerThrIleTyrAsnAsnGluValLeuHis---GlnProLeuSerGluAla 176
   |||
   |||
   |||
Db 2383 GAT-----GACCTAAAGCTAACAGCTGTCCTTCGTTCTCAGCAACTGCTAGAACT 2433
   |||
   |||
   |||
QY 177 GlnArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArg--- 195
   |||
   |||
   |||
Db 2434 CAAAGTAACTTGGAAAGCCTTAAGCCTTAAGTTCCAACACACAGAGAGATCTCTTAC-AGA 2492
   |||
   |||
   |||
QY 196 -----GluAsnThrLeuValSerGluHisAlaGlnArgAspGlnArgGluThrGlnCys 213
   |||
   |||
   |||
Db 2493 AAAGCAGAAAGAGGTTGTTTCAGATCCATCGGAGAGAGACTAGCCCAACCCAGTGT 2552
   |||
   |||
   |||
QY 214 GlnMetLysGluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThr 233
   |||
   |||
   |||
Db 2553 CAAATCCAGGAAATTCACAAATGTATCAAAATAAACAATGTAACCTGAAGGAATTCATTG 2612
   |||
   |||
   |||
QY 234 GluGlnGlnGluSerLeuAspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeu 253
   |||
   |||
   |||
Db 2613 GAAAG-CAGGAATCTGTAGAGCGGCGGTGTCCTCAACTCCAGAGTGAATAATGTGCTT 2671
   |||
   |||
   |||
QY 254 GlnGlnGlnLeuValHisAlaHisLysLysAlaAspAsnLysSerLysIleThrIleAsp 273
   |||
   |||
   |||
Db 2672 CCACAGCAACTGGATCATCTTACAAACAAAGCTGACAAATCAAGAAAGACAAATATTACT 2731
   |||
   |||
   |||
QY 274 IleHisPheLeuGluArgLysMetGlnHisHisLeu-----LeuLys 287
   |||
   |||
   |||
```

```
Db 2732 ATCCAAGACCAATCTCGTCTACTGTGAAAAATCTTCAAGCTGAGAAAGAAAGCTAGAA 2791
QY 288 GluLysAsnGluGluIlePheAsnTyrAsnAsnHisLeuLysAsnArgIleTyrGlnTyr 307
   |||
   |||
   |||
Db 2792 GAGAAAATCAGACATTATCAACGAATGCAATCAATTTTTTTTAAAGGCTTCAACACAG 2851
   |||
   |||
   |||
QY 308 GluLysGluLysAlaGluThrGlu 315
   |||
   |||
   |||
Db 2852 GGAAATGAGAAAGCAGAAAGAGAA 2875
   |||
   |||
   |||
RESULT 7
LOCUS BU678550/c
DEFINITION BU678550 638 bp mRNA linear EST 07-OCT-2002
          UI-CF-DUI-aar-o-16-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone
          UI-CF-DUI-aar-o-16-0-UI 3', mRNA sequence.
ACCESSION BU678550
VERSION BU678550.1 GI:23525612
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 638)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
       discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
          University of Iowa
          2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
          Tel: 319 356 4866
          Fax: 319 356 7171
          Email: paul-mccray@uiowa.edu
          Tissue Procurement: Dr. M. J. Welsh, University of Iowa
          cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
          cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
          DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Researchers may obtain clones from Research
          Genetics (www.regen.com) or from Open Biosystems
          (www.openbiosystems.com).
          Seq primer: M13 FORWARD
          POLYA=Yes.
FEATURES             Location/Qualifiers
     1..638
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="UI-CF-DUI-aar-o-16-0-UI"
     /tissue_type="Primary Lung Epithelial Cells"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
     /clone_lib="UI-CF-DUI"
     /note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
     modified polylinker; Site 1: EcoR I; Site 2: Not I;
     UI-CF-DUI is a normalized cDNA library containing the
     following tissue(s): Primary Lung Epithelial Cells The
     library was constructed according to Bonaldo, Lennon and
     Soares, Genome Research, 6:791-806, 1996. First strand
     cDNA synthesis was primed with an oligo-dT primer
     containing a Not I site. Double stranded cDNA was ligated
     to an EcoR I adaptor, digested with Not I, and cloned
     directionally into pT7T3-Pac vector. The oligonucleotide
     used to prime the synthesis of first-strand cDNA contains
     a library tag sequence that is located between the Not I
     site and the (dT)18 tail. The sequence tag for this
     library is GGCTGTAGGC.
     TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
     TAG_LIB=UI-CF-DUI
     TAG_SEQ=GGCTGTAGGC"
```


COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgaaps-r@mail.nih.gov
 Tissue Procurement: CLONTECH Laboratories, Inc.
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLC1360 row: d column: 11
 High quality sequence stop: 730.
 Location/Qualifiers

FEATURES

source

1..862
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4612186"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 75"
 /notes="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.65
 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 7,51e-41 Length: 862
 Score: 481.50 Matches: 112
 Percent Similarity: 64.32% Conservative: 34
 Best Local Similarity: 49.34% Mismatches: 76
 Query Match: 29.50% Indels: 5
 DB: 4 Gaps: 2
 US-09-489-079-25 (1-317) x BG431116 (1-862)

QY 1 MetGlyThrArgAlaLeuGlnCysGluValSerHisThrHisGluAsnGluAenTyrLeu 20
 DB 146 ATGGCTCAAAAGAAATGAATTCCTGAGAATTCCTCATGATGACGAGCAAAAGACCTA 205
 QY 21 LeuHisGluAsnCysMetLeuLysGluLeuAlaMetLeuLysGluLeuAlaThr 40
 DB 206 TCGCATAAATAATAGCATGTTGCGAGGAGAAATTCCTATGCTAAGACTAGAAATAGACACA 265
 QY 41 LeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysLeuLys 60
 DB 266 ATAAAAATCAAAACAGGAAAGAAAGAAAGAAATGTTCTGAGGACCTTAAATGTATATA 325
 QY 61 GluLysAsnAlaGluLeuGlnMetThrLeuLysGluLysGluSerLeuThrLysArg 80
 DB 326 GAAAGAATCAAGACCTTCAGAGACTATATAACAGAAATGAGGAAACATTAACACAAACA 385
 QY 81 AlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThrSer 100
 DB 386 ATATCCAGTATAATGACGGCTAGTGTCTGACAGCTGAGAAATCAATGCTAAATCTT 445
 QY 101 LysLeuLys--GluLysGlnAspLysGluLeuLeuGluAlaGluLeuGluSerHisHis 119
 DB 446 AAATGAGAAATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 505
 QY 120 ProArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSerGln 139
 DB 506 TCTAGATGGCTGCTGCTATACATGATCGTGATCAAAAGTCAGACATCAAAAAAGAGAAC 565
 QY 140 --GluProAlaPheHisIleAlaGlyAspAlaCys-----LeuGlnArgLysMetAsnV 157
 DB 566 TAGAACTGCTTCCAGAGAGCAAGAGATGAATGTTCTCTGTTTACAGGACAAATGAATT 625

QY 157 alAspValSerSerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAlaG 177
 DB 626 TTGATGTGTCTAACTAAAGATAACATGAGATTCTTCTCAACAATATTTAAACTGA 685
 QY 177 lnaArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGlu 197
 DB 686 AAGTAAACTCAATAGTCTTAGAAATTTAGTTTCATTCACAGAGAGATGCGCTCAGAGAAC 745
 QY 197 snThrLeuValSerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysG 217
 DB 746 AGACTTGGGGCTTAGAACGGGTACCAAGGGCCCTAGCCAAACACACAGGTCACATGAAG 805
 QY 217 luAlaGluHisMetTyr 222
 DB 806 CAATGGACCAAAACGTTT 822
 RESULT 11
 BP238820
 LOCUS
 DEFINITION
 BP238820 Sugano cDNA library, liver HepG2 Homo sapiens cDNA clone
 HEP04550, mRNA sequence.
 ACCESSION
 BP238820
 VERSION
 BP238820.1 GI:52111730
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 600)
 AUTHORS
 Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 JOURNAL
 Contact: Yutaka Suzuki
 COMMENT
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.
 FEATURES
 Location/Qualifiers
 1..600
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEP04550"
 /tissue_type="liver"
 /cell_line="HepG2"
 /clone_lib="Sugano cDNA library, liver HepG2"
 /note="hepatoma"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,64e-40 Length: 600
 Score: 470.00 Matches: 107
 Percent Similarity: 68.00% Conservative: 29
 Best Local Similarity: 53.50% Mismatches: 61
 Query Match: 28.80% Indels: 4
 DB: 5 Gaps: 2
 US-09-489-079-25 (1-317) x BP238820 (1-600)
 QY 40 ThrLeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspIleLysLeu 59
 DB 2 TCAATAAAAAATCAAAACCCAGGAAAGAAAGAAAGAAATGTTTTCAGGACCTTAAAT--TTG 59
 QY 60 LysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLys 79
 DB 60 AAAAGAAATTTATGAAGACCTTGAAGAGACTATATAACAGAAATGAGGAAACATTAACACAA 119
 QY 80 ArgAlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThr 99
 DB 120 ACAATATCCAGTATAATGACGGCTAGTGTCTGACAGCTGAGAAATGCAATGCTAAT 179

LOCUS BQ926802 1014 bp mRNA linear EST 20-AUG-2002
 DEFINITION AGENCOURT 8795642 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6429928
 5', mRNA sequence.
 ACCESSION BQ926802
 VERSION BQ926802.1 GI:22341833
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1014)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-b@mail.nih.gov
 Tissue Procurement: DCTD/DTP/Gazdar
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LFCM2618 row: o column: 17
 High quality sequence start: 24
 High quality sequence stop: 682.
 Location/Qualifiers
 1..1014
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6429928"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_18"
 /notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

FEATURES

source

Alignment Scores:
 Pred. No.: 6,57e-34 Length: 1014
 Score: 417.50 Matches: 117
 Percent Similarity: 60.15% Conservative: 46
 Best Local Similarity: 43.17% Mismatches: 91
 Query Match: 25.58% Indels: 20
 DB: Gaps: 5

US-09-489-079-25 (1-317) x BQ926802 (1-1014)

QY 1 MetGlyThr-ArgAlaLeuGlnCysGluValSer---HisThrHisGluAsnGluAsnTy 19
 Db 792 ATGGGAACCAAGAAATTCGTCAGTTTCGATGTTCTGATGTCAGGAAAGAGGAGGC 733
 QY 19 rLeuLeuHisGluAsnCysMetLeuLysGluLeuAlaMetLeuLysLeuGluLeuAl 39
 Db 732 CCTGTTGCATATAAAACCATTTCGATGCAAGATGAAATTCGCCAGGCTCAGGCTGGAAATACA 673
 QY 39 aThrLeuLysHisGlnTyrglnGlnGlyGluAsnLysTyrrPheGluAsnLysLeuLe 59
 Db 672 CACAATAAAATCAAACTCTCGAAAG-----AAATACCTTAAAGACATTGAAATAT 619
 QY 59 uLysGluLysAsnAlaGluLeuGlnMetThrLeuLysLeuLysGluGluSerLeuThrLy 79
 Db 618 AAAAGAAACATGACGACTTTCAAAGGCTTAAACAGAAATGGGAAATCAACAAA 559
 QY 79 sArgAlaSerGlnTyrsrGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuTh 99

Db 558 AACGATAGCCCATTTATAGTGGACAGCTTACTGCTCTGACAGATGAGAACACATGCTCCG 499
 QY 99 rSerLysLeuLysGluLysGlnAspLysGluLeuLeuGluAlaGluLeuGluSerHisH 119
 Db 498 TTCTAACTGGAGAGGAAAAACA-AAGCAGGCAAGAGCTGACGAAATGGAATCATACCA 440
 QY 119 sProArgLeuAlaSerAlaValGlnAspHisGlnIleValThrSerArgLysSerG 139
 Db 439 TCGTAGACTGAATGTTGCTCTCATGATCATCATCAAGTCACTCAACAGAAAGAGACCA 380
 QY 139 nGluProAlaPheHisIleAlaGlyAspAlaCys-----LeuGlnArgLysMetAsnVa 157
 Db 379 AGAGCTTGATTTCCAGGCGACAGTAGATAAATTCGCTCATTTACAGAAAATTTGAATTC 320
 QY 157 lAspValSerSerThrIleTyrrAsnAsnGluValLeuHisGlnProLeuSerGluAlaG 177
 Db 319 TCATGTTCTG-----ATTCCTTCTCAGCAACTTTCTAAAGCTGA 281
 QY 177 nArgLysSerLysSerLeuLysIleAsnLeuAsnTyrrAlaGlyAspAlaLeuArgL 197
 Db 280 GAGTAGTCCAGGTGCTCCAAACTGAGCTCCATACCAAGAGAGGCTCTGAAAGAAA 221
 QY 197 nThrLeuValSerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLysG 217
 Db 220 CGCTTTGGTTTTGAACATACATACCAAGAGAGCTTAACCAACGG--TGTCAAATGAAGGA 163
 QY 217 uAlaGluHisMetTyrrGlnAsnGlnGlnAsnValAsnLysHisThrGlnGlnGln 237
 Db 162 CATTGAAAAATGTACACAAAGTGGCCCAAGATAAGTAGAAAATATGCAAGAAAGCAA-- 105
 QY 237 uSerLeuAspGlnLysLeuPheGlnLeuGlnSerLysAsnMetTrpLeuGlnGlnLe 257
 Db 104 -----GTAAGATTATGTCACTACGAAAGACAAAATATGTTGCTTCAACAGCAACT 55
 QY 257 uValHisAlaHisLysLysAlaAspAsnLys 267
 Db 54 GGATGATT-CACAAAGGCTTCGTCG 25

RESULT 14

LOCUS CN415409 670 bp mRNA linear EST 16-MAY-2004
 DEFINITION 17000532179892 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN415409
 VERSION CN415409.1 GI:47403003
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 670)
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
 Lebkowski, J. and Stanton, L.W.

TITLE

Transcriptome characterization elucidates signaling networks that
 control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 CONTACT: Brandenberger R
 Regenerative Medicine
 Genon Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@genon.com
 Insert Length: 670 Std Error: 0.00.
 Location/Qualifiers
 1..670
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /issue_type="embryonic stem cells, cell lines H1, H7, and
 H9"
 /clone_lib="GRN_ES"
 /notes="oligo dT primed, full-length enriched cDNA library"

FEATURES

source

from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

```

ORIGIN
Alignment Scores:
Pred. No.: 5,3e-34 Length: 670
Score: 416.00 Matches: 99
Percent Similarity: 61.54% Conservative: 45
Best Local Similarity: 42.31% Mismatches: 74
Query Match: 25.49% Indels: 16
DB: 7 Gaps: 5

US-09-489-079-25 (1-317) x CN415409 (1-670)

QY 20 LeuLeuHisGluAsnCysMetLeuLysGluLeuAlaMetLeuLysLeuGluLeuAla 39
DB 8 CTGTCGATGAACCGCTTGATCAAGATGAATGCGCAGCTCAGGCTGGAAGAAGAC 67
QY 40 ThrLeuLysHisGlnTyrGlnGluLysGluAsnLysTyrPheGluAspLeuLysLeu 59
DB 68 ACATATAAAACCAAACTGGAAG-----AAATCTTAAAGACTTGAATTTG 121
QY 60 LysGluLysAsnAlaGluLeuMetThrLeuLysLeuLysGluGluSerLeuThrLys 79
DB 122 AAAAGAAAGCATGAAGACCTTCAAAGGCTCTAAACGGAATGGGGAACATTAGCAAAA 181
QY 80 ArgAlaSerGlnTyrSerGlyGlnLeuLysValLeuLeuAlaGluAsnThrMetLeuThr 99
DB 182 ACGATAGCTGTATATAGTGACAGCTGTCTCTGACAGATGAAACACACACGCTCCGT 241
QY 100 SerLysLeu---LysGluLysGlnAspLysGluLeuGluAlaGluLeuSerHis 118
DB 242 TCATAACTGGAGACCAAGACAGCAGGCAAGACTGGAACAGAAATGCATCATAC 301
QY 119 HisProArgLeuAlaSerAlaValGlnAspHisAspGlnIleValThrSerArgLysSer 138
DB 302 CATTGTAGACTGAATGCTCGATGTGATCATGATCAAAAGTCACCTCAAAAAGAGAC 361
QY 139 GlnGluProAlaPheHisIleAlaGlyAspAlaCys-----LeuGlnArgLysMetAsn 156
DB 362 CAAGAGCTTGCTTCCAGGGCCACAGTAGATAAATGCTGCTACTTACAGGAAATTTGAAT 421
QY 157 ValAspValSerSerThrIleTyrAsnAsnGluValLeuHisGlnProLeuSerGluAla 176
DB 422 TCTCATGTTCTG-----ATCTTCTCTGCAACTTCTTAAAGCT 460
QY 177 GlnArgLysSerLysSerLeuLysIleAsnLeuAsnTyrAlaGlyAspAlaLeuArgGlu 196
DB 461 GAGAGTAAGTCCAGAGTCTCTCAAACTGAGCTCCATTAACAGGAGAGGCTCTGAAAGAA 520
QY 197 AsnThrLeuValSerGluHisAlaGlnArgAspGlnArgGluThrGlnCysGlnMetLys 216
DB 521 AAGGCTTTGGTTTGAACACAGTGCMAAGTGAGCTAAGCAAAACACAGAGTCAAAATGAAG 580
QY 217 GluAlaGluHisMetTyrGlnAsnGluGlnAspAsnValAsnLysHisThrGlnGlnGln 236
DB 581 GACATTGAAAAATGTACAAAAGTGGATACATACATGGAATGGAATGCGTAGAAAAACAG 640
QY 237 GluSerLeuAspGlnLysLeuPheGlnLeuGlnSerLysAsn 250
DB 641 GAA-----AGATTTTGTCACTTAAAAAACAATAAT 670

```

```

RESULT 15
BQ437838/c
LOCUS
DEFINITION
AGENCOURT 7892743 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6157635
5', mRNA sequence.
ACCESSION
BQ437838
VERSION
BQ437838.1 GI:21176914
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 886)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13503 Row: n column: 04
High quality sequence stop: 682.
Location/Qualifiers
1. 886
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6157635"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."

FEATURES

source

ORIGIN

```

Alignment Scores:
Pred. No.: 6,37e-33 Length: 886
Score: 407.50 Matches: 93
Percent Similarity: 57.59% Conservative: 55
Best Local Similarity: 36.19% Mismatches: 108
Query Match: 24.97% Indels: 1
DB: 5 Gaps: 1

US-09-489-079-25 (1-317) x BQ437838 (1-886)

QY 12 HisThrHisGluAsnGluAsnTyrIleLeuHisGluAsnCysMetLeuLysGluIle 31
DB 775 CATATCATGAAGAAATGAAAGATCTGATGATTAATAATTCATTNTGATATACATATT 716
QY 32 AlaMetLeuLysLeuGluIleAlaThrLeuLysHisGlnTyrGlnGluLysGluAsnLys 51
DB 715 GCTATATTTCAGACAGGAAATGTGCAATGAAAAATGACACCTGGAAAAAGAAATAAA 656
QY 52 TyrPheGluAspIleLysIleLeuLysGluLysAsnAlaGluLeuGlnMetThrLeuLys 71
DB 655 TAACTTAAGGAAGTTAAATTTGTAAGAAACAATGCTGCTTGGAAAGTATATAAAA 596
QY 72 LeuLysGluGluSerLeuThrLysArgAlaSerGlnTyrSerGlyGlnLeuLysValLeu 91
DB 595 CTCATGAGGAATTCATAACAATAACAGCATTCGCGTATCAACAGAGCTTAATGATCTC 536
QY 92 IleAlaGluAsnThrMetLeuThrSer---LysLeuLysGluLysGlnAspLysGluIle 110
DB 535 AAAGCTGAGATAACAAGGCTCAGTTCCGGACTGTTGAAGGAAGAAAGCAACAAAGA 476
QY 111 LeuGluAlaGluIleGluSerHisHisProArgLeuAlaSerAlaValGlnAspHis 130
DB 475 CTGGAAGCTGAAACTGAATCTTATCAGTCTAGACTGGCTGCTTCTATAGTAACACAGT 416
QY 131 GlnIleValThrSerArgLysSerGlnGluProAlaPheHisIleAlaGlyAspAlaCys 150
DB 415 GAAAGTGTGAAAAACAGAAAGAAACCTAAACCTTTAGAGAGAACTCAAGATGTTTCC 356
QY 151 LeuGlnArgLysMetAsnValAspValSerSerThrIleTyrAsnAsnGluValLeuHis 170
DB 355 GTACAAAGTAAATAATGAGTTCTGATATTTCCGAAAGTAGAAGATAGAGTATGTTTCTTACT 296

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:04:01 ; Search time 15.6398 Seconds
(without alignments)
1088.086 Million cell updates/sec

Title: US-09-489-079-26

Perfect score: 225

Sequence: 1 MQKSVNKALELNQETLRA.....PSESKQDYESSWDSISLC 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1990s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	44	3 AAB07640	Aab07640 Amino aci
2	225	100.0	168	4 AAB63909	Aab63909 Human pro
3	225	100.0	169	4 AAB63906	Aab63906 Human pro
4	225	100.0	174	4 AAB63926	Aab63926 Human pro
5	225	100.0	185	4 AAB63933	Aab63933 Human pro
6	225	100.0	189	4 AAB63929	Aab63929 Human pro
7	225	100.0	195	4 AAB63918	Aab63918 Human pro
8	225	100.0	207	4 AAB63937	Aab63937 Human pro
9	225	100.0	220	4 AAB63917	Aab63917 Human pro
10	225	100.0	223	4 AAB63903	Aab63903 Human pro
11	225	100.0	225	4 AAB63901	Aab63901 Human pro
12	225	100.0	229	4 AAB63925	Aab63925 Human pro
13	225	100.0	241	7 ADB83983	Adb83983 Human bre
14	225	100.0	266	4 AAB63899	Aab63899 Human pro
15	225	100.0	398	3 AAB07638	Aab07638 Amino aci
16	225	100.0	445	4 AAB50249	Aab50249 Human bre
17	225	100.0	445	4 AAG65987	Aag65987 B726P spl
18	225	100.0	445	4 AAU33350	Aau33350 Human bre
19	225	100.0	445	5 AAG78917	Aag78917 Human bre
20	225	100.0	445	6 ABJ37740	Abj37740 Human tum
21	225	100.0	445	7 ADL93135	Adl93135 Human bre
22	225	100.0	445	8 ADE44425	Ade44425 Human bre
23	225	100.0	466	4 AAB50248	Aab50248 Human bre
24	225	100.0	466	4 AAG65986	Aag65986 B726P spl
25	225	100.0	466	4 AAU33349	Aau33349 Human bre

ALIGNMENTS

RESULT 1

AAB07640

ID AAB07640 standard; peptide; 44 AA.

XX

AC AAB07640;

XX

DT 07-NOV-2000 (first entry)

XX

DE Amino acid sequence of an antigenic peptide from BS322 polypeptide.

XX

KW BS322; breast tissue marker; breast disease; breast cancer.

XX

OS Homo sapiens.

XX

PN WO200043420-A1.

XX

PD 27-JUL-2000.

XX

PF 21-JAN-2000; 2000WO-US001452.

XX

PR 21-JAN-1999; 99US-00234716.

XX

PA (ABBO) ABBOTT LAB.

XX

PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;

PI Stroupe SD;

XX

DR WPI; 2000-499217/44.

XX

PT BS322-specific polynucleotides, polypeptides and antibodies, used in the detection and diagnosis of breast disease especially breast cancer.

XX

PS Claim 23; Page 125; 126pp; English.

XX

CC The present sequence represents an antigenic peptide derived from human BS322 polypeptide. The peptide is used to raise antibodies. BS322 is a breast tissue marker. The BS322 polynucleotides and polypeptides are used to detect and diagnose breast disease, e.g. breast cancer. The BS322 polynucleotides are useful as a source of probes and primers, and the BS322 polypeptides are useful as antigens

XX

SQ Sequence 44 AA;

Query Match 100.0%; Score 225; DB 3; Length 44;

Best Local Similarity 100.0%; Pred. No. 9.4e-23;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MOKSPVKALELKNQTLRADEILPSESKQDYESSWDSLSL 44
 |||
 Db 1 MOKSPVKALELKNQTLRADEILPSESKQDYESSWDSLSL 44

RESULT 2

ID AAB63909 standard; protein; 168 AA.
 XX
 AC AAB63909;
 DT 26-MAR-2001 (first entry)
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1271.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 XX cancer associated antigen; cytostatic; cancer vaccine.
 KW Homo sapiens.
 XX WO200073801-A2.
 XX 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US014749.
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX Example 1; Page 772; 799pp; English.

Query Match 100.0%; Score 225; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 5.4e-22;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MOKSPVKALELKNQTLRADEILPSESKQDYESSWDSLSL 44
 |||
 Db 107 MOKSPVKALELKNQTLRADEILPSESKQDYESSWDSLSL 150

RESULT 3

ID AAB63906 standard; protein; 169 AA.
 XX
 AC AAB63906;
 DT 26-MAR-2001 (first entry)
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1268.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 XX WO200073801-A2.
 XX 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US014749.
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 XX WO200073801-A2.
 XX 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US014749.
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX Example 1; Page 770-771; 799pp; English.

AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX Sequence 169 AA;

Query Match 100.0%; Score 225; DB 4; Length 169;
 Best Local Similarity 100.0%; Pred. No. 5.4e-22;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MOKSPVKALELKNQTLRADEILPSESKQDYESSWDSLSL 44
 |||
 Db 91 MOKSPVKALELKNQTLRADEILPSESKQDYESSWDSLSL 134

RESULT 4

ID AAB63926 standard; protein; 174 AA.
 XX
 AC AAB63926;
 DT 26-MAR-2001 (first entry)
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1288.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 XX WO200073801-A2.
 XX 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US014749.
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Example 1; Page 779; 799pp; English.
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 174 AA;
 SQ Query Match 100.0%; Score 225; DB 4; Length 174;
 Best Local Similarity 100.0%; Pred. No. 5.6e-22;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOKSPVKALELKNQTLRADEILPESKQDYESSWDSLSLC 44
 DB 19 MOKSPVKALELKNQTLRADEILPESKQDYESSWDSLSLC 62
 RESULT 5
 AAB63933
 ID AAB63933 standard; protein; 185 AA.
 XX
 AC AAB63933;
 XX
 DT 26-MAR-2001 (first entry)
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1295.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 XX cancer associated antigen; cytostatic; cancer vaccine.
 KW Homo sapiens.
 XX
 OS WO200073801-A2.
 XX
 PN 07-DEC-2000.
 XX
 PD 26-MAY-2000; 2000WO-US014749.
 XX
 PF 28-MAY-1999; 99US-0136526P.
 XX
 PR 10-SEP-1999; 99US-0153454P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Example 1; Page 782; 799pp; English.
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 185 AA;
 SQ Query Match 100.0%; Score 225; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 6.1e-22;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOKSPVKALELKNQTLRADEILPESKQDYESSWDSLSLC 44
 DB 90 MOKSPVKALELKNQTLRADEILPESKQDYESSWDSLSLC 133
 RESULT 6
 AAB63929
 ID AAB63929 standard; protein; 189 AA.
 XX
 AC AAB63929;
 XX
 DT 26-MAR-2001 (first entry)
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1291.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 XX
 PR 10-SEP-1999; 99US-0153454P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Example 1; Page 780; 799pp; English.
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 189 AA;
 SQ Query Match 100.0%; Score 225; DB 4; Length 189;

CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 185 AA;
 SQ Query Match 100.0%; Score 225; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 6.1e-22;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MOKSPVKALELKNQTLRADEILPESKQDYESSWDSLSLC 44
 DB 90 MOKSPVKALELKNQTLRADEILPESKQDYESSWDSLSLC 133
 RESULT 6
 AAB63929
 ID AAB63929 standard; protein; 189 AA.
 XX
 AC AAB63929;
 XX
 DT 26-MAR-2001 (first entry)
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1291.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 XX
 PR 10-SEP-1999; 99US-0153454P.
 XX
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Example 1; Page 780; 799pp; English.
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 189 AA;
 SQ Query Match 100.0%; Score 225; DB 4; Length 189;

```
Best Local Similarity 100.0%; Pred. No. 6.3e-22; Mismatches 0; Gaps 0;
Matches 44; Conservative 0; Indels 0;

QY 1 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 44
    |||||
Db 112 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 155

RESULT 7
AAB63918
ID AAB63918 standard; protein; 195 AA.
XX
AC AAB63918;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1280.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
XX 10-SEP-1999; 99US-0153454P.
XX (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PS Nucleic acids encoding breast, gastric and prostate cancer associated
XX antigen precursors, useful for diagnosing and treating a condition
XX characterized by expression of an abnormal amount of a protein, e.g.
XX cancer.
XX
PS Example 1; Page 775-776; 799pp; English.
XX
CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
XX represent nucleotide sequences encoding human breast, gastric and
XX prostate cancer associated antigen precursors (CAAP) respectively.
XX AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
XX represent human breast, gastric and prostate CAAP protein sequence
XX respectively. CAAPs have cytostatic activity and can be used in the
XX production of cancer vaccines. The human CAAP proteins, peptides, nucleic
XX acids or anti-CAAP antibodies are useful for diagnosing and treating a
XX condition characterised by expression of an abnormal amount of a protein,
XX e.g. cancer.
XX
PS Sequence 195 AA;
XX
QY Query Match 100.0%; Score 225; DB 4; Length 195;
XX Best Local Similarity 100.0%; Pred. No. 6.5e-22;
XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 44
    |||||
Db 78 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 121

RESULT 8
AAB63937
ID AAB63937 standard; protein; 207 AA.
XX
AC AAB63937;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1279.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PS Sequence 207 AA;
XX
QY Query Match 100.0%; Score 225; DB 4; Length 207;
XX Best Local Similarity 100.0%; Pred. No. 7.1e-22;
XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 44
    |||||
Db 11 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 54

RESULT 9
AAB63917
ID AAB63917 standard; protein; 220 AA.
XX
AC AAB63917;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1279.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PS Sequence 207 AA;
XX
QY Query Match 100.0%; Score 225; DB 4; Length 207;
XX Best Local Similarity 100.0%; Pred. No. 7.1e-22;
XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 44
    |||||
Db 11 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 54

RESULT 9
AAB63917
ID AAB63917 standard; protein; 220 AA.
XX
AC AAB63917;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1279.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PS Sequence 207 AA;
XX
QY Query Match 100.0%; Score 225; DB 4; Length 207;
XX Best Local Similarity 100.0%; Pred. No. 7.1e-22;
XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 44
    |||||
Db 11 MQKSPNKALELKNQETLRADILPSESKQDYESSWDSLSL 54
```



```

XX 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
PI
XX Obata Y;
XX WPI; 2001-025274/03.
XX
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
XX Example 1; Page 775; 799pp; English.
XX
XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX
XX Sequence 220 AA;
SQ
Query Match 100.0%; Score 225; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 7.7e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQKSVNKALELNQETLRADILPSESQKDYESSWDSLSIC 44
Db 8 MQKSVNKALELNQETLRADILPSESQKDYESSWDSLSIC 51
XX
RESULT 10
AAB63903
ID AAB63903 standard; protein; 223 AA.
XX
XX AAB63903;
XX
XX 26-MAR-2001 (first entry)
XX
XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1265.
XX
XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
XX cancer associated antigen; cytostatic; cancer vaccine.
XX
XX Homo sapiens.
XX
XX WO200073801-A2.
XX
XX 07-DEC-2000.
XX
XX 26-MAY-2000; 2000WO-US014749.
XX
XX 28-MAY-1999; 99US-0136526P.
XX
XX 10-SEP-1999; 99US-0153454P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Obata Y;
XX
XX WPI; 2001-025274/03.
XX
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX

```

```

XX Example 1; Page 769; 799pp; English.
XX
XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX
XX Sequence 223 AA;
SQ
Query Match 100.0%; Score 225; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 7.8e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQKSVNKALELNQETLRADILPSESQKDYESSWDSLSIC 44
Db 98 MQKSVNKALELNQETLRADILPSESQKDYESSWDSLSIC 141
XX
RESULT 11
AAB63901
ID AAB63901 standard; protein; 225 AA.
XX
XX AAB63901;
XX
XX 26-MAR-2001 (first entry)
XX
XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1263.
XX
XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
XX cancer associated antigen; cytostatic; cancer vaccine.
XX
XX Homo sapiens.
XX
XX WO200073801-A2.
XX
XX 07-DEC-2000.
XX
XX 26-MAY-2000; 2000WO-US014749.
XX
XX 28-MAY-1999; 99US-0136526P.
XX
XX 10-SEP-1999; 99US-0153454P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Obata Y;
XX
XX WPI; 2001-025274/03.
XX
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
XX Example 1; Page 768; 799pp; English.
XX
XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX

```

```

SQ Sequence 225 AA;
Query Match          100.0%; Score 225; DB 4; Length 225;
Best Local Similarity 100.0%; Pred. No. 7.9e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKSVPNKALELKNEQTLRADEILPSESQKQDYESSWDSLSLC 44
   |||||
DB 16 MOKSVPNKALELKNEQTLRADEILPSESQKQDYESSWDSLSLC 59
   |||||

RESULT 12
AAB63925
ID AAB63925 standard; protein; 229 AA.
XX
AC AAB63925;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1287.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
PS Example 1; Page 778-779; 799pp; English.
XX
CC AAP22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX
SQ Sequence 229 AA;
Query Match          100.0%; Score 225; DB 4; Length 229;
Best Local Similarity 100.0%; Pred. No. 8.1e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKSVPNKALELKNEQTLRADEILPSESQKQDYESSWDSLSLC 44
   |||||
DB 19 MOKSVPNKALELKNEQTLRADEILPSESQKQDYESSWDSLSLC 62
   |||||

RESULT 13
ADB83983
ID ADB83983 standard; protein; 241 AA.

```

```

XX ADB83983;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human breast cancer diagnostic marker protein Incyte 411152.
XX
KW cancer; neurodegenerative disorder; human; breast cancer;
KW diagnostic marker.
XX
OS Homo sapiens.
XX
FH Key
FT Misc-difference 72 Location/Qualifiers
FT /note= "Encoded by GAG"
FT Misc-difference 75 /note= "Encoded by GAT"
FT Misc-difference 76 /note= "Encoded by AAT"
FT Misc-difference 79 /note= "Encoded by TTT"
FT Misc-difference 83 /note= "Encoded by CCC"
FT Misc-difference 85 /note= "Encoded by AGA"
XX
PN US2003104418-A1.
XX
PD 05-JUN-2003.
XX
PF 25-APR-2002; 2002US-00133757.
XX
PR 27-APR-2001; 2001US-0287153P.
XX
PA (ZHAN/) ZHANG C.
PA (MAHI/) MAHINI B.
PA (WALK/) WALKER M G.
XX
PI Zhang C, Mahini B, Walker MG;
XX
DR WPI; 2003-687833/65.
DR N-PSDB; ADB83986.
XX
PT New combination of polynucleotides, useful for preparing a composition
PT for diagnosing or treating cancer or neurodegenerative disorders.
XX
PS Example 11; Page 20-21; 22pp; English.
XX
CC The invention related to a combination of polynucleotides. The
CC combination of polynucleotides is useful for preparing a composition for
CC diagnosing or treating cancer or neurodegenerative disorders. The present
CC sequence represents the amino acid sequence of the human breast cancer
CC diagnostic marker Incyte 411152.
XX
SQ Sequence 241 AA;
Query Match          100.0%; Score 225; DB 7; Length 241;
Best Local Similarity 100.0%; Pred. No. 8.6e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOKSVPNKALELKNEQTLRADEILPSESQKQDYESSWDSLSLC 44
   |||||
DB 1 MOKSVPNKALELKNEQTLRADEILPSESQKQDYESSWDSLSLC 44
   |||||

RESULT 14
AAB63899
ID AAB63899 standard; protein; 266 AA.
XX
AC AAB63899;
XX
DT 26-MAR-2001 (first entry)
XX

```

DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1261.

XX Human: breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.

XX Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014749.

XX 28-MAY-1999; 99US-0136526P.

XX 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.

XX Example 1; Page 767; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer

XX Sequence 266 AA;

Query Match 100.0%; Score 225; DB 4; Length 266;
Best Local Similarity 100.0%; Pred. No. 9.8e-22;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQKSVNKALELNQTLRADEILPSESQKDYESSWDSLSLC 44
DB 22 MQKSVNKALELNQTLRADEILPSESQKDYESSWDSLSLC 65

RESULT 15

AAB07638

ID AAB07638 standard; protein; 398 AA.

XX AC AAB07638;

XX DT 07-NOV-2000 (first entry)

XX Amino acid sequence of BS322 polypeptide.

XX BS322; breast tissue marker; breast disease; breast cancer.

XX Homo sapiens.

XX WO200043420-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-US001452.

XX 21-JAN-1999; 99US-00234716.

XX

PA (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
PI Stroupe SD;

XX WPI; 2000-499217/44.

XX DR N-PSDB; AAA59015.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
PT detection and diagnosis of breast disease especially breast cancer.

XX Claim 23; Page 122-123; 126pp; English.

XX The present sequence represents a human BS322 polypeptide. BS322 is a
CC breast tissue marker. The BS322 polynucleotides and polypeptides are used
CC to detect and diagnose breast disease, e.g. breast cancer. The BS322
CC polynucleotides are useful as a source of probes and primers, and the
CC BS322 polypeptides are useful as antigens

XX Sequence 398 AA;

Query Match 100.0%; Score 225; DB 3; Length 398;
Best Local Similarity 100.0%; Pred. No. 1.7e-21;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQKSVNKALELNQTLRADEILPSESQKDYESSWDSLSLC 44
DB 1 MQKSVNKALELNQTLRADEILPSESQKDYESSWDSLSLC 44

Search completed: August 1, 2005, 22:18:04
Job time : 15.6398 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:08:06 ; Search time 2.7109 Seconds
(without alignments)
1561.672 Million cell updates/sec

Title: US-09-489-079-26
Perfect score: 225
Sequence: 1 MOKSVNKALEKNEQTILRA.....PSESKQDYESSWDSLSLC 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	61	27.1	629	T15945	hypothetical prote
2	60.5	26.9	1203	B55094	chromosomal protei
3	59.5	26.4	678	T32483	hypothetical prote
4	59	26.2	771	T38616	hypothetical prote
5	59	26.2	938	T39006	related to yeast z
6	58	25.8	257	G90727	molybdate-binding
7	58	25.8	257	H85578	molybdate-binding
8	57.5	25.6	552	T04653	receiver-like prot
9	57	25.3	876	G90592	hypothetical prote
10	57	25.3	1015	JC6552	DNA topoisomerase
11	56.5	25.1	707	F83465	flagellar biosynth
12	56	24.9	173	T47279	hypothetical prote
13	56	24.9	752	S51866	HPRI protein - yea
14	55.5	24.7	237	A85901	probable alpha hel
15	55.5	24.7	237	A49940	probable alpha hel
16	55.5	24.7	237	E91056	endospore developm
17	55.5	24.7	332	C40646	hypothetical prote
18	55.5	24.7	554	S46775	molybdate-binding
19	55	24.4	257	C64812	transmembrane prot
20	55	24.4	491	AH1140	transmembrane prot
21	55	24.4	491	AC1499	hypothetical prote
22	55	24.4	616	B85508	hypothetical prote
23	55	24.4	616	E90657	hypothetical prote
24	54.5	24.2	274	T40098	hypothetical prote
25	54.5	24.2	1240	T48800	SMT4 related prote
26	54	24.0	396	D64509	hypothetical prote
27	54	24.0	666	C84861	hypothetical prote
28	54	24.0	713	S46838	hypothetical prote
29	54	24.0	1104	C72409	reverse gyrase - T

RESULT 1
TI5945
hypothetical protein F01F1.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15945
R:Miller, N.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid F01F1.
A:Reference number: Z18435
A:Accession: T15945
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-629 <MIL>
A:Cross-references: UNIPROT:Q19086; EMBL:U13070; NID:G529697; PID:G529702; PIDN:AAC4663
A:Introns: 120/2; 286/3; 384/1; 478/2
C:Genetics:
A:Experimental source: strain Bristol N2
C:Superfamily: Synechocystis hypothetical protein slr0825

ALIGNMENTS

Query Match 27.1%; Score 61; DB 2; Length 629;
Best Local Similarity 32.1%; Pred. No. 19;
Matches 18; Conservative 5; Mismatches 15; Indels 18; Gaps 2;
Qy 1 MOKSVN-----KALEKNEQTLRADEILPSESKQDYESSWDSLSL 43
Db 184 MQSLPMPWDETSTIRMAIDKGGSSNEVTLK-----DGTGKQINYSPTWDGDEL 234
RESULT 2
B55094
chromosomal protein XCAP-E - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: B55094
R:Hirano, T.; Mitchison, T.J.
Cell 79, 449-458, 1994
A:Title: A heterodimeric coiled-coil protein required for mitotic chromosome condensation
A:Reference number: A55094; MUID:95042742; PMID:7954811
A:Accession: B55094
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1203 <HIR>
A:Cross-references: UNIPROT:P50533; GB:U13674; NID:G563813; PIDN:AAA64680.1; PID:G56381
C:Superfamily: chromosome segregation protein SMCI
C:Keywords: chromosomal protein; DNA condensation; heterodimer
Query Match 26.9%; Score 60.5; DB 2; Length 1203;
Best Local Similarity 40.5%; Pred. No. 46;
Matches 15; Conservative 8; Mismatches 13; Indels 1; Gaps 1;


```

Qy      1 MQSKVPKALELNKEQTILRADEL---PSESKQKDY---EESWDS 40
           | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      87 MDYAVDKAIDTATROTLGNLSLVVAVPKASEQKDFTIDSKTNWTS 132

RESULT 8
T04653
receiver-like protein 4 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04653; T52035
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 1999
A;Reference number: Z15263
A;Accession: T04653
A:Molecule type: DNA
A;Residues: 1-552 <BEV>
A;Cross-references: UNIPROT:O49397; EMBL:AL021636
A;Experimental source: cultivar Columbia; BAC clone F10N7
R;Buchholz, G.; Keitel, C.; Kircher, S.; Kudla, J.; Lohrmann, J.; Sweere, U.; Schaefer,
submitted to the EMBL Data Library, May 1998
A;Description: Nuclear-localized receiver-like proteins are differentially expressed in
A;Reference number: Z25912
A;Accession: T52035
A;Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: mRNA
A;Residues: 1-552 <BUC>
A;Cross-references: EMBL:AJ005195; PIDN:CAA06432.1
C;Genetics:
A;Gene: ARP4
A;Map position: 4
A;Intons: 43/1; 94/1; 207/1; 232/3; 517/3
A;Note: F10N7.270
C;Superfamily: response regulator, ARR1 type; response regulator homology
P;19-129/Domain: response regulator homology <RRH>

Query Match          25.6%; Score 57.5; DB 2; Length 552;
Best Local Similarity 30.4%; Pred.No. 43;
Matches 17; Conservative 9; Mismatches 15; Indels 15; Gaps 3;

Qy      2 QKSVPNKALELNKEQTILRA-----DEILPS-----ESKQKYEESSWDSLESL 43
           | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      480 RSSNPNEG-LMGQQKLQSGLMASDAGSLDDIVNSLTQTQSOSDFSEGWDLDGL 534

RESULT 9
G90592
hypothetical protein MYPU 6470 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: G90592
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, P.; Moszer, I.; A.
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: G90592
A;Status: preliminary
A:Molecule type: DNA
A;Residues: 1-876 <CUR>
A;Cross-references: UNIPROT:Q9BP82; GB:AL445566; PID:g14090062; PIDN:CAC13820.1; GSPDB:G
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU 6470
A;Genetic code: SGC3

Query Match          25.3%; Score 57; DB 2; Length 876;
Best Local Similarity 36.8%; Pred.No. 84;
Matches 14; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

Qy      5 VPNKALELNKEQTILRADELPSKQKYEESSWDSLESL 42
           | : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      562 LPNNLLAKNVKMLTVDEIFRLNLKNLPENTKDONS 599

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:05:16 ; Search time 12.3033 Seconds
(without alignments)
1831.334 Million cell updates/sec

Title: US-09-489-079-26
Perfect score: 225
Sequence: 1 MQKSVNKALELNQTLRA.....PSESKQKQYESSWDSLSL 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	180	80.0	1011	2 Q9BXX2	Q9bxx2 homo sapien
3	65	28.9	823	2 Q9H0H6	Q9h0h6 homo sapien
4	62	27.6	223	2 Q9QZM8	Q9qzm8 rattus norv
5	62	27.6	851	2 Q8IIR2	Q8iir2 plasmodium
6	61	27.1	226	2 Q9D5A2	Q9d5a2 mus musculu
7	61	27.1	234	2 Q80ZP0	Q80zpo mus musculu
8	61	27.1	629	2 Q19086	Q19086 caenorhabdi
9	60.5	26.9	899	2 Q9SG75	Q9sg75 arabidopsis
10	60.5	26.9	917	2 Q9CAP6	Q9caf6 arabidopsis
11	60.5	26.9	1203	1 SMC2_XENLA	P50533 xenopus lae
12	60	26.7	391	2 Q8KRC9	Q8krc9 myxococcu
13	59.5	26.4	204	2 Q6W3L3	Q6w3l3 alvinella p
14	59.5	26.4	678	2 Q17399	Q17399 caenorhabdi
15	59	26.2	257	2 Q8CW92	Q8cw92 escherichia
16	59	26.2	380	2 Q6LXV6	Q6lxv6 methanococ
17	59	26.2	684	2 Q803Y7	Q803y7 brachydanio
18	59	26.2	785	2 Q8V5S9	Q8v5s9 helicoverpa
19	59	26.2	785	2 Q1BWS5	Q1bws5 helicoverpa
20	59	26.2	785	2 Q99G11	Q99g11 helicoverpa
21	59	26.2	938	2 Q14100	Q14100 schizosacch
22	59	26.2	2228	2 Q8I1W5	Q8i1w5 plasmodium
23	58.5	26.0	336	2 Q7QC17	Q7qc17 anopheles g
24	58.5	26.0	348	2 Q86JW4	Q86jw4 dictyosteli
25	58.5	26.0	938	2 Q66I11	Q66i11 xenopus tro
26	58.5	26.0	1319	2 Q9HCD3	Q9hcd3 homo sapien
27	58.5	26.0	1504	2 Q8IKW4	Q8ikw4 plasmodium
28	58.5	26.0	4524	2 Q8I3J9	Q8i3j9 plasmodium
29	58	25.8	257	2 Q8X9J1	Q8x9j1 escherichia
30	58	25.8	257	2 Q83SA5	Q83sa5 shigella fl
31	58	25.8	294	2 Q9NW22	Q9nw22 homo sapien

32	58	25.8	548	2	Q7XPN5	Q7xpn5 oryza sativ
33	58	25.8	585	2	Q9W4C4	Q9w4c4 drosophila
34	58	25.8	607	2	Q7KVM1	Q7kvm1 drosophila
35	58	25.8	706	2	Q96HM2	Q96hm2 homo sapien
36	58	25.8	706	2	Q9BTX2	Q9btx2 homo sapien
37	58	25.8	706	2	Q9NVM5	Q9nvm5 homo sapien
38	58	25.8	706	2	Q9NVM9	Q9nvm9 homo sapien
39	58	25.8	1066	2	Q7XE90	Q7xe90 oryza sativ
40	58	25.8	1654	2	Q7RHE8	Q7rhe8 plasmodium
41	58	25.8	1793	1	YCF1_LOTJA	Q9bbn6 lotus japon
42	57.5	25.6	134	2	Q7P5L6	Q7p5l6 fusobacteri
43	57.5	25.6	209	2	Q7Z2X1	Q7z2x1 homo sapien
44	57.5	25.6	246	2	Q7PXD8	Q7pxd8 anopheles g
45	57.5	25.6	272	2	Q9IL99	Q9il99 white spot

ALIGNMENTS

RESULT 1
Q9BXX3 PRELIMINARY; PRT; 1341 AA.
AC Q9BXX3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Breast cancer antigen NY-BR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-Specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269087; AAK27325.1; -;
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00023; Ank; 6
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
KW ANK repeat.
SQ SEQUENCE 1341 AA; 152776 MW; 33B53DDE6FD3A58B CRC64;

Query Match 100.0%; Score 225; DB 2; Length 1341;
Best Local Similarity 100.0%; Pred. No. 8.5e-17;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQKSVNKALELNQTLRADEILPSESKQKQYESSWDSLSL 44
Db 711 MQKSVNKALELNQTLRADEILPSESKQKQYESSWDSLSL 754

RESULT 2
Q9BXX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BXX2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guerre A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AP269088; AAK27326.1; -.
DR HSSP; O75832; IUOH.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DE438 CRC64;

Query Match 80.0%; Score 180; DB 2; Length 1011;
Best Local Similarity 79.5%; Pred. No. 8.5e-12;
Matches 35; Conservative 4; Mismatches 5; Indels 0; Gaps

QY 1 MQKSVNKALEKNEOTLRADILPSESKQKDYESSWDSLSLC 44
DB 385 MQKTVNKAFLKNEQTLRAAQMFSESKQKDESNWDSSEPC 428
|||||
|||||

RESULT 3
Q9H0H6 PRELIMINARY; PRT; 823 AA.
AC Q9H0H6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFp434A171.
GN Name=DKFp434A171;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mawes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.,
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136793; CAB66727.1; -.
DR HSSP; Q60778; IOY3.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKVRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 823 AA; 93968 MW; 9D2DDDSF52C379B2 CRC64;

```

Query Match 28.9%; Score 65; DB 2; Length 823;
Best Local Similarity 50.0%; Pred. No. 89;
Matches 15; Conservative 7; Mismatches 6; Indels 2; Gaps 1;

QY	9	A	E	L	K	N	E	Q	T	R	A	D	E	I	L	P	-	S	E	S	K	Q	D	Y	E	S	36
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:		
Db	349	A	V	Q	R	K	N	V	Q	T	R	A	E	Q	A	L	P	V	A	S	E	E	E	E	E	E	378

RESULT 4
Q9QZW8

ID	Q9QZW8	PRELIMINARY;	PRT;	223 AA.
AC	Q9QZW8;			
DT	01-MAY-2000 (TrEMBLrel. 13, Created)			
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE	Dnaj-like protein.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Rattus.			
OC	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Wistar; TISSUE=Testis;			
RL	Wang L., Miao S., Yang J., Zhang X., Zhao M., Li M.;			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF154849; AAD53061.1; -.			
DR	HSSP; P25685; 1HDJ.			
DR	InterPro; IPR001623; DnaJ_N.			
DR	Pfam; PF00226; DnaJ; 1.			
DR	SMART; SM00271; DnaJ; 1.			
DR	PROSITE; PS00636; DNAJ_1; 1.			
DR	PROSITE; PS00706; DNAJ_2; 1.			
DR	SEQUENCE 223 AA; 25756 MW; EBB4ADFC3CC92D3F CRC64;			
QY	8 KALELKEQLRADEILPSKSKDKYEESSW 38			
DB	40 EAAREKFKVAEAYQLSDAKKRDYDRSW 70			
RESULT 5				
Q8II2	PRELIMINARY;	PRT;	851 AA.	
ID	Q8II2	PRELIMINARY;	PRT;	851 AA.
AC	Q8II2;			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	Hypothetical protein.			
GN	ORFNames=PF11_0232;			
OS	Plasmodium falciparum (isolate 3D7).			
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.			
OC	NCBI_TaxID=36329;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22255705; PubMed=12368864; DOI=10.1038/nature01097;			
RA	Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,			
RA	Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,			
RA	Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,			
RA	Chan M.S., Nene V., Shallow J., Suh B., Peterson J., Angiuoli S.,			
RA	Chen T.M., Allen J., Selengut J., Haft D., Macher M.W., Vaidya A.B.,			
RA	Martin D.M., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,			
RA	McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,			
RA	Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,			
RA	Fraser C.M., Barrell B.;			
RT	"Genome sequence of the human malaria parasite Plasmodium falciparum";			
RL	Nature 419:498-511(2002).			
RL	EMBL; AE014839; AAN35816.1; -.			
SW	Hypothetical protein.			
QY	SEQUENCE 851 AA; 101937 MW; 466174BC463FD823 CRC64;			

Query Match 27.6%; Score 62; DB 2; Length 851;
Best Local Similarity 35.1%; Pred. No. 2e+02;
Matches 13; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

Qy 3 KSPVKNAL~~E~~LKNEQTTRADEILPSEKQKYESSWD 39
| | | | : : | : | : | : | : | :
Dd 5 KONEKKGLTNETIIEKENKEVEEKEKEFESEYN 41

DR SMART; SM00434; TOP4c; 1.
 DR TIGRFPMS; TIGR01063; GYR2A; 1.
 SQ SEQUENCE 917 AA; 101412 MW; 7568C4004D524976 CRC64;
 Query Match 26.9%; Score 60.5; DB 2; Length 917;
 Best Local Similarity 43.8%; Pred. No. 3.3e+02;
 Matches 14; Conservative 8; Mismatches 9; Indels 1; Gaps 1;
 QY 9 ALEKNEQTLRADEILPSESQKDYESSWDS 40
 DB 785 AMELKNEKDIASNDIIPA-SLRKMEKSEDA 815
 RESULT 11
 SMC2_XENLA
 ID SMC2_XENLA STANDARD; PRT; 1203 AA.
 AC P50533;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Structural maintenance of chromosome 2 (Chromosome-associated protein
 DE E) (Chromosome assembly protein XCAP-E).
 GN .Name=SMC2; Synonym=XCAP-E;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A., AND INTERACTION WITH XCAP-C.
 RX MEDLINE=95042742; PubMed=7954811; DOI=10.1016/S0092-8674(94)90254-2;
 RA Hirano T., Mitchison T.J.;
 RT "A heterodimeric coiled-coil protein required for mitotic chromosome
 RT condensation in vitro";
 RL Cell 79:449-458 (1994).
 RN [2]
 RP IDENTIFICATION IN A CONDENSIN COMPLEX WITH XCAP-C; XCAP-H; XCAP-D2 AND
 RP XCAP-G.
 RX MEDLINE=97304380; PubMed=9160743; DOI=10.1016/S0092-8674(00)80233-0;
 RA Hirano T., Kobayashi R., Hirano M.;
 RT "Condensins, chromosome condensation protein complexes containing
 RT XCAP-C, XCAP-E and a Xenopus homolog of the Drosophila Barren
 RT protein";
 RL Cell 89:511-521 (1997).
 RN [3]
 RP FUNCTION OF THE CONDENSIN COMPLEX.
 RX MEDLINE=98447791; PubMed=9774278; DOI=10.1126/science.282.5388.487;
 RA Kimura K., Hirano M., Kobayashi R., Hirano T.;
 RT "Phosphorylation and activation of 13S condensin by Cdc2 in vitro.";
 RL Science 282:487-490 (1998).
 RN [4]
 RP FUNCTION OF THE CONDENSIN COMPLEX.
 RX MEDLINE=99354989; PubMed=10428035; DOI=10.1016/S0092-8674(00)81018-1;
 RA Kimura K., Rybenkov V.V., Crisano N.J., Hirano T., Cozzarelli N.R.;
 RT "13S condensin actively reconfigures DNA by introducing global
 RT positive writhe: implications for chromosome condensation";
 RL Cell 98:239-248 (1999).
 RN [5]
 RP FUNCTION: Central component of the condensin complex, a complex
 RP required for conversion of interphase chromatin into mitotic-like
 RP condense chromosomes. The condensin complex probably introduces
 RP positive supercoils into relaxed DNA in the presence of type I
 RP topoisomerases and converts nicked DNA into positive knotted forms
 RP in the presence of type II topoisomerases.
 CC -1- SUBUNIT: Forms a heterodimer with XCAP-C/SMC4L1. Component of the
 CC condensin complex, which contains the XCAP-E/SMC2L1 and XCAP-
 CC C/SMC4L1 heterodimer, and three non SMC subunits that probably
 CC regulate the complex: XCAP-H/BRN1, XCAP-D2/CNAP1 and XCAP-G/CAPG.
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. In interphase
 CC cells, the majority of the condensin complex is found in the
 CC cytoplasm, while a minority of the complex is associated with
 CC chromatin. A subpopulation of the complex however remains
 CC associated with chromosome foci in interphase cells. During
 CC mitosis, most of the condensin complex is associated with the

chromatin. At the onset of prophase, the regulatory subunits of
 the complex are phosphorylated by Cdc2, leading to condensin's
 association with chromosome arms and to chromosome condensation.
 Dissociation from chromosomes is observed in late telophase (By
 similarity).
 -1- DOMAIN: The hinge domain, which separates the large intramolecular
 coiled coil regions, allows the heterodimerization with XCAP-C,
 forming a V-shaped heterodimer (By similarity).
 -1- SIMILARITY: Belongs to the SMC family. SMC2 subfamily.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 EMBL; UI3674; AAA64680.1; --
 FIR; B55094; B55094.
 DR HSSP; P04268; IIC2.
 DR InterPro; IPR003405; SMC C.
 DR InterPro; IPR010935; SMC hinge.
 DR InterPro; IPR003395; SMC_N.
 DR Pfam; PF02483; SMC_C; 1.
 DR Pfam; PF06470; SMC hinge; 1.
 DR Pfam; PF02463; SMC_N; 1.
 DR ATP-binding; Cell Cycle; Coiled coil; DNA condensation; Mitosis;
 KW Nuclear protein.
 FT NP_BIND 32 39 ATP (Potential).
 FT DOMAIN 172 513 Coiled coil (Potential).
 FT DOMAIN 514 669 Flexible hinge.
 FT DOMAIN 670 1032 Coiled coil (Potential).
 FT DOMAIN 1095 1121 Ala/Asp-rich (DA-box).
 SQ SEQUENCE 1203 AA; 136341 MW; 04323DD0027DF309 CRC64;
 Query Match 26.9%; Score 60.5; DB 1; Length 1203;
 Best Local Similarity 40.5%; Pred. No. 4.4e+02;
 Matches 15; Conservative 8; Mismatches 13; Indels 1; Gaps 1;
 QY 2 QKSVNPKALEKNEQTLRADEILPSESQKDYESSW 38
 DB 182 QKTIKKEAKLKEIQILSEETPTIKLKE-ERSY 217
 RESULT 12
 O8KRC9
 ID O8KRC9 PRELIMINARY; PRT; 391 AA.
 AC O8KRC9;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE DnaJ.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophactetaceae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=DZfl;
 RA Ueki T., Inoue S.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Acts as a co-chaperone. Stimulates, jointly with grpE,
 CC the ATPase activity of dnaK (By similarity).
 CC -1- COFACTOR: Binds 2 zinc ions per monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the dnaJ family.
 DR EMBL; AY033943; AAK59395.1; --
 DR HSSP; P25685; 1HDJ.
 DR GO; GO:0051082; F:unfolded protein binding; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR GO; GO:0006457; P:protein folding; IEA.
 DR GO; GO:0006986; P:response to unfolded protein; IEA.

```

DR InterPro; IPR002939; DnaJ C.
DR InterPro; IPR001305; DnaJ_CXXCXGK.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR008971; HSP40_DnaJ_pap.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXXCXGK; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS0076; DnaJ_2; 1.
KW Chapterone; DNA replication; Heat shock; Metal-binding; Repeat; Zinc.
SQ SEQUENCE 391 AA; 40914 MW; 195A823B8618BFCE CRC64;

Query Match 26.7%; Score 60; DB 2; Length 391;
Best Local Similarity 40.5%; Pred. No. 1.4e+02;
Matches 15; Conservative 6; Mismatches 16; Indels 0; Gaps 0;

QY 7 NKALELKNQOTLRADILPSESQKDYESSWDSLSL 43
DB 39 NKAABEKQVSAFEVLSDTRRKLDYDFGPDAEKI 75

RESULT 13
Q6W3L3 PRELIMINARY; PRT; 204 AA.
AC Q6W3L3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lipoprotein.
GN Name=NT03AF0015;
OS Alvinella pompejana epibiont 703.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae.
OX NCBI_TaxID=244800;
RV [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22837667; PubMed=12957888;
RA DOI=10.1128/AEM.69.9.5070-5078.2003;
RA Campbell B.J., Stein J.L., Cary S.C.;
RT "Evidence of chemolithoautotrophy in the bacterial community
RT associated with Alvinella pompejana, a hydrothermal vent polychaete.";
RL Appl. Environ. Microbiol. 69:5070-5078(2003).
DR EMBL; AY12991; AAQ75168.1; -.
KW Lipoprotein.
SQ SEQUENCE 204 AA; 24167 MW; 8617AEFB26A1F13F CRC64;

Query Match 26.4%; Score 59.5; DB 2; Length 204;
Best Local Similarity 41.9%; Pred. No. 81;
Matches 18; Conservative 8; Mismatches 10; Indels 7; Gaps 3;

QY 1 MOKSVNKALELKNQOTLRADILPSES--KQKD-----YEES 37
DB 20 MKLKVPNKVLENIIDT-KSNILPLEGYKLGKGNFLYSENS 61

RESULT 14
O17399 PRELIMINARY; PRT; 678 AA.
AC O17399;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein F52H2.1.
GN Name=F52H2.1; ORFNames=F52H2.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RV [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Chisoe S., Hawkins J.;
RT "The sequence of C. elegans cosmid F52H2.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF026214; AAB71311.1; -.
DR PIR; T32483; T32483.
DR WormBase; WBGene00018714; F52H2.1.
DR WormPep; F52H2.1a; CE10864.
KW Hypothetical protein.
SQ SEQUENCE 678 AA; 78262 MW; 49ABB658D5919B6 CRC64;

Query Match 26.4%; Score 59.5; DB 2; Length 678;
Best Local Similarity 42.9%; Pred. No. 3e+02;
Matches 15; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

QY 2 QKSVNKALELKNQOTLRADILPSESQKDYESS 36
DB 619 KXNKKNAKELQEQPLR-HQFRPTKSFNDYLS 652

RESULT 15
Q8CW92 PRELIMINARY; PRT; 257 AA.
AC Q8CW92;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Molybdate-binding periplasmic protein.
GN Name=modA; OrderedlocusNames=c0840;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RV [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.F., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR EMBL; AE016757; AAN79313.1; -.
DR HSP; P37329; 1W0D.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR GO; GO:0015412; F:molybdate-transporting ATPase activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0015689; P:molybdate ion transport; IEA.

```

DR GO:0006810; P:transport; IEA.
DR InterPro; IPR011587; MBP prokaryotic.
DR InterPro; IPR005950; MolP ABC transpt.
DR InterPro; IPR006059; SBP Bac_I.
DR InterPro; IPR010916; TONB_Box_N.
DR Pfam; PF01547; SBP_Bac_I; 1.
DR ProDom; PD008688; MBP prokaryotic; 1.
DR TIGRFAMs; TIGR01256; modA; 1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 257 AA; 27414 MW; CECFE98ED902AE87 CRC64;

Query Match 26.2%; Score 59; DB 2; Length 257;
Best Local Similarity 30.4%; Pred. NO. 1.2e+02;
Matches 14; Conservative 12; Mismatches 14; Indels 6; Gaps 2;

QY 1 MQKSVPNKALELKNQETLRADEL---PSESQKDY--EESWDS 40
DB 87 MDYAVDKKAITASRTLLGNLSLVVAVPKASEQKQDFTIDSKTNWTS 132

Search completed: August 1, 2005, 22:22:18
Job time : 16.3033 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:24:58 ; Search time 666.152 Seconds
(without alignments)
3200.517 Million cell updates/sec

Title: US-09-489-079-26

Perfect score: 225
Sequence: 1 MQKSPNKALELKNQTLRA.....PSEKQKQYESSWDSISLC 44

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09489079/runat_01082005_141250_13297/app_query.fasta_1.1635
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1 1 8858 @runat_01082005_141250_13297 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.cm.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	278	6	BD271305
2	225	100.0	445	6	CQ419026 Sequence
3	225	100.0	533	6	CQ423647 Sequence
4	225	100.0	562	6	CQ432495 Sequence

C	5	225	100.0	572	6	CQ427744	CQ427744 Sequence
	6	225	100.0	578	6	CQ427463	CQ427463 Sequence
C	7	225	100.0	580	6	AX156306	AX156306 Sequence
	8	225	100.0	636	6	AX053794	AX053794 Sequence
	9	225	100.0	664	6	AX053793	AX053793 Sequence
	10	225	100.0	708	6	AX053785	AX053785 Sequence
	11	225	100.0	713	6	CQ417876	CQ417876 Sequence
	12	225	100.0	713	6	AX053787	AX053787 Sequence
	13	225	100.0	741	6	AX053804	AX053804 Sequence
	14	225	100.0	766	6	AX053799	AX053799 Sequence
	15	225	100.0	770	6	CQ427927	CQ427927 Sequence
	16	225	100.0	771	6	AX053802	AX053802 Sequence
	17	225	100.0	776	6	AX053782	AX053782 Sequence
	18	225	100.0	814	6	AX053780	AX053780 Sequence
	19	225	100.0	833	6	AX053800	AX053800 Sequence
	20	225	100.0	904	6	AX053778	AX053778 Sequence
	21	225	100.0	1158	6	AX053798	AX053798 Sequence
	22	225	100.0	1337	6	AR283455	AR283455 Sequence
	23	225	100.0	1337	6	AR344223	AR344223 Sequence
	24	225	100.0	1337	6	AR351424	AR351424 Sequence
	25	225	100.0	1337	6	AR454004	AR454004 Sequence
	26	225	100.0	1337	6	AR561592	AR561592 Sequence
	27	225	100.0	1337	6	AX282974	AX282974 Sequence
	28	225	100.0	1337	6	AX303147	AX303147 Sequence
	29	225	100.0	1448	6	CQ490684	CQ490684 Sequence
	30	225	100.0	1448	6	CQ495497	CQ495497 Sequence
	31	225	100.0	1448	6	CQ496522	CQ496522 Sequence
	32	225	100.0	1729	6	AR283454	AR283454 Sequence
	33	225	100.0	1729	6	AR344222	AR344222 Sequence
	34	225	100.0	1729	6	AR351423	AR351423 Sequence
	35	225	100.0	1729	6	AR454003	AR454003 Sequence
	36	225	100.0	1729	6	AR561591	AR561591 Sequence
	37	225	100.0	1729	6	AX282973	AX282973 Sequence
	38	225	100.0	1729	6	AX303146	AX303146 Sequence
	39	225	100.0	2030	6	AX367056	AX367056 Sequence
	40	225	100.0	2232	6	AR283468	AR283468 Sequence
	41	225	100.0	2232	6	AR454017	AR454017 Sequence
	42	225	100.0	2232	6	AX303171	AX303171 Sequence
	43	225	100.0	2307	6	AR283456	AR283456 Sequence
	44	225	100.0	2307	6	AR344224	AR344224 Sequence
	45	225	100.0	2307	6	AR351425	AR351425 Sequence

ALIGNMENTS

RESULT 1
BD271305
LOCUS BD271305 278 bp DNA linear PAT 17-JUL-2003
DEFINITION Reagents and methods useful for detecting diseases of the breast.
ACCESSION BD271305
VERSION BD271305.1 GI:33081073
KEYWORDS JP 2002540761-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 278)
AUTHORS Medel, P.A.B., Cohen, M., Colpitts, T.L., Friedman, P.N., Gordon, J., Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D., Russell, J.C., and Scroupe, S.D.
TITLE Reagents and methods useful for detecting diseases of the breast
JOURNAL Patent: JP 2002540761-A 2 03-DEC-2002;
ABBOTT LABORATORIES
COMMENT OS Homo sapiens (human)
PN JP 2002540761-A/2
PD 03-DEC-2002
PF 21-JAN-2000 JP 2000594836
PR 21-JAN-1999 US 09/234716
PI PATRICIA A BILLING MEDEL, MAURICE COHEN, TRACEY L COLPITTS, PAULA N FRIEDMAN, EDWARD N GRANADOS, STEVEN C HODGES, MICHAEL R PI JULIAN GORDON, PI

PI JON D KRATOCHVIL, JOHN C RUSSELL, STEPHEN D STROUPE PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68, G01N33/53, G01N33/53, G01N33/566, G01N33/574,
PC G01N37/00,
PC C12N15/00, C12N5/00
CC Reagents and methods useful for detecting diseases of the CC
breast

PH Key Location/Qualifiers
FT source 1. .278
/organism='Homo sapiens (human)'.
FEATURES
source
1. .278
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 9.36e-25 Length: 278
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x BD271305 (1-278)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
|||||
Db 4 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGAATTTGAAGATGAACAACATTGAGAGCA 63
|||||

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
64 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTTGGGATTCT 123
|||||

QY 41 GluSerLeuCys 44
124 GAGAGTCTCTGT 135
|||||

RESULT 2
CQ419026
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .445
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 1.56e-24 Length: 445
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x CQ419026 (1-445)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
|||||
Db 67 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGAATTTGAAGATGAACAACATTGAGAGCA 126
|||||

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
127 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTTGGGATTCT 186
|||||

QY 41 GluSerLeuCys 44
187 GAGAGTCTCTGT 198
|||||

RESULT 4
CQ432495
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
|||||
Db 39 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGAATTTGAAGATGAACAACATTGAGAGCA 98
|||||

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
99 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTTGGGATTCT 158
|||||

QY 41 GluSerLeuCys 44
159 GAGAGTCTCTGT 170
|||||

RESULT 3
CQ423647
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .533
/organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 1.9e-24 Length: 533
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x CQ423647 (1-533)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
67 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGAATTTGAAGATGAACAACATTGAGAGCA 126
|||||

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
127 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTTGGGATTCT 186
|||||

QY 41 GluSerLeuCys 44
187 GAGAGTCTCTGT 198
|||||

RESULT 4
CQ432495
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

assessment, prevention, and therapy of breast cancer

Patent: WO 0151628-A 17529 19-JUL-2001;

Millennium Pharmaceuticals, Inc. (US)

FEATURES

source

1..562

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,02e-24 Length: 562

Score: 225.00 Matches: 44

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x CQ432495 (1-562)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20

Db 45 ATGCAAAAGTCTGTTCCAAATAAGCTTGGATTGAAGATGAACAACATTGAGAGCA 104

Qy 21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40

Db 105 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 164

Qy 41 GluSerLeuCys 44

Db 165 GAGAGTCTCTGT 176

RESULT 5

LOCUS

CQ427744/c Sequence 12778 from Patent WO0151628.

DEFINITION

CQ427744

ACCESSION

VERSION

CQ427744.1 GI:41379973

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.

Novel genes, compositions, kits, and methods for identification,

assessment, prevention, and therapy of breast cancer

Patent: WO 0151628-A 12778 19-JUL-2001;

Millennium Pharmaceuticals, Inc. (US)

FEATURES

source

1..572

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,05e-24 Length: 572

Score: 225.00 Matches: 44

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x CQ427744 (1-572)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20

Db 528 ATGCAAAAGTCTGTTCCAAATAAGCTTGGATTGAAGATGAACAACATTGAGAGCA 469

Qy 21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40

Db 468 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 409

Qy 41 GluSerLeuCys 44

Db 408 GAGAGTCTCTGT 397

RESULT 6

LOCUS

CQ427463 Sequence 12497 from Patent WO0151628.

DEFINITION

CQ427463

ACCESSION

VERSION

CQ427463.1 GI:41379692

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.

Novel genes, compositions, kits, and methods for identification,

assessment, prevention, and therapy of breast cancer

Patent: WO 0151628-A 12497 19-JUL-2001;

Millennium Pharmaceuticals, Inc. (US)

FEATURES

source

1..578

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,08e-24 Length: 578

Score: 225.00 Matches: 44

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x CQ427463 (1-578)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20

Db 57 ATGCAAAAGTCTGTTCCAAATAAGCTTGGATTGAAGATGAACAACATTGAGAGCA 116

Qy 21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40

Db 117 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 176

Qy 41 GluSerLeuCys 44

Db 177 GAGAGCCTCTGT 188

RESULT 7

LOCUS

AX156306/c Sequence 163 from Patent WO0140269.

DEFINITION

AX156306

ACCESSION

VERSION

AX156306.1 GI:14537307

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Dillon, D.C., Day, C.H., Jiang, Y., Houghton, R.L., Mitcham, J.L. and

Wang, A.

Compositions and methods for therapy and diagnosis of breast cancer

Patent: WO 0140269-A 163 07-JUN-2001;

CORIXA CORPORATION (US)

FEATURES

source

1..580

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

```
ORIGIN
Alignment Scores:
Pred. No.: 2,09e-24 Length: 580
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x AX156306 (1-580)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAla 20
Db 523 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGATTGAAGATGAACAACATTGAGAGCA 464
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 463 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTCTTTGGGATTCT 404
QY 41 GluSerLeuCys 44
Db 403 GAGAGTCTCTGT 392

RESULT 8
AX053794
LOCUS AX053794 636 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 560 from Patent WO0073801.
ACCESSION AX053794
VERSION AX053794.1 GI:12228113
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Obata,Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor
JOURNAL Patent: WO 0073801-A 560 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
1..636
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2,31e-24 Length: 636
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x AX053794 (1-636)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAla 20
Db 234 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGATTGAAGATGAACAACATTGAGAGCA 293
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 294 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTCTTTGGGATTCT 353
QY 41 GluSerLeuCys 44
Db 354 GAGAGTCTCTGT 365

RESULT 9
AX053793
LOCUS AX053793 664 bp DNA linear PAT 13-JAN-2001
```

```
DEFINITION Sequence 559 from Patent WO0073801.
ACCESSION AX053793
VERSION AX053793.1 GI:12228112
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Obata,Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor
JOURNAL Patent: WO 0073801-A 559 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
1..664
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2,42e-24 Length: 664
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x AX053793 (1-664)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLysAsnGluGlnThrLeuArgAla 20
Db 24 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGATTGAAGATGAACAACATTGAGAGCA 83
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 84 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTCTTTGGGATTCT 143
QY 41 GluSerLeuCys 44
Db 144 GAGAGTCTCTGT 155

RESULT 10
AX053785
LOCUS AX053785 708 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 551 from Patent WO0073801.
ACCESSION AX053785
VERSION AX053785.1 GI:12228104
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Obata,Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor
JOURNAL Patent: WO 0073801-A 551 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
1..708
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2,59e-24 Length: 708
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
```

```
DB: 6 Gaps: 0
US-09-489-079-26 (1-44) x AX053785 (1-708)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
DB 271 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 330
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
DB 331 GATGAGATACTCCATCAGATCCAAACAAAGGACTATGAGAAAGTCTTGGATTCT 390
QY 41 GluSerLeuCys 44
DB 391 GAGAGTCTCTGT 402
RESULT 11
LOCUS CO417876 713 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 2910 from Patent WO0151628.
ACCESSION CO417876
VERSION CO417876.1 GI:41370105
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
AUTHORS Novel genes, compositions, kits, and methods for identification,
TITLE assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 2910 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .713
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 2.61e-24 Length: 713
Pred. No.: 225.00 Matches: 44
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-489-079-26 (1-44) x CQ417876 (1-713)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
DB 100 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 159
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
DB 160 GATGAGATACTCCATCAGATCCAAACAAAGGACTATGAGAAAGTCTTGGATTCT 219
QY 41 GluSerLeuCys 44
DB 220 GAGAGCTCTGT 231
RESULT 12
LOCUS AX053787 713 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 553 from Patent WO0073801.
ACCESSION AX053787
VERSION AX053787.1 GI:12228106
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE
```

```
REFERENCE 1
AUTHORS Obata, Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor
JOURNAL Patent: WO 0073801-A 553 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
1. .713
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 2.61e-24 Length: 713
Pred. No.: 225.00 Matches: 44
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-489-079-26 (1-44) x AX053787 (1-713)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
DB 319 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 378
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
DB 379 GATGAGATACTCCATCAGATCCAAACAAAGGACTATGAGAAAGTCTTGGATTCT 438
QY 41 GluSerLeuCys 44
DB 439 GAGAGTCTCTGT 450
RESULT 13
LOCUS AX053804 741 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 570 from Patent WO0073801.
ACCESSION AX053804
VERSION AX053804.1 GI:12228123
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Obata, Y.
JOURNAL Breast, gastric and prostate cancer associated antigens and uses
therefor
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
1. .741
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 2.73e-24 Length: 741
Pred. No.: 225.00 Matches: 44
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-09-489-079-26 (1-44) x AX053804 (1-741)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
DB 32 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 91
```

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 92 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTGGGATTCT 151

Qy 41 GluSerLeuCys 44
Db 152 GAGAGTCTCTGT 163

RESULT 14
AX053799 LOCUS 766 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 565 from Patent WO0073801.
ACCESSION AX053799
VERSION AX053799.1 GI:12228118

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Obata, Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor

JOURNAL Patent: WO 0073801-A 565 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)

FEATURES
source
1. .766
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2.83e-24 Length: 766
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x AX053799 (1-766)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 262 ATGCAAAAGTCTGTTCCAAATTAAGCCTTGGAATTGAAGATGAACAACATTTGAGAGCA 321

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 322 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTGGGATTCT 381

Qy 41 GluSerLeuCys 44
Db 382 GAGAGTCTCTGT 393

RESULT 15
CQ427927 LOCUS 770 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 12961 from Patent WO0151628.
ACCESSION CQ427927
VERSION CQ427927.1 GI:41380156

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Lillie, J., Xu, Y., Wang, Y., and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 12961 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source
1. .770
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2.84e-24 Length: 770
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-26 (1-44) x CQ427927 (1-770)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 39 ATGCAAAAGTCTGTTCCAAATTAAGCCTTGGAATTGAAGATGAACAACATTTGAGAGCA 98

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 99 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTGGGATTCT 158

Qy 41 GluSerLeuCys 44
Db 159 GAGAGTCTCTGT 170

Search completed: August 2, 2005, 02:51:20
Job time : 669.152 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:23:28 ; Search time 81.4834 Seconds
(without alignments)
3196.585 Million cell updates/sec

Title: US-09-489-079-26

Perfect score: 225

Sequence: 1 MOKSVNKALELNEQTLRA.....PSESKQDYESSWDSLELC 44

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPIO.spool_p/US09489079/runat_01082005_141249_13290/app_query.fasta_1.1635
-DB=N Geneseq 16Dec04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1 1 1418 @runat_01082005_141249_13290 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	278	3 AAAS9008	Aaa59008 Nucleotid
2	225	100.0	445	4 AAL11595	Aal11595 Human bre
3	225	100.0	533	4 AAL16216	Aal16216 Human bre
4	225	100.0	562	4 AAL25059	Aal25059 Human bre
5	225	100.0	572	4 AAL20310	Aal20310 Human bre

6	225	100.0	578	4 AAL20029	Aal20029 Human bre
7	225	100.0	580	4 AAH55638	Aah55638 Human bre
8	225	100.0	580	12 ADN40408	Adn40408 Human bre
9	225	100.0	600	11 ACN92774	Acn92774 Breast ca
10	225	100.0	635	11 ACN87508	Acn87508 Breast ca
11	225	100.0	636	4 AAF22981	Aaf22981 Human pro
12	225	100.0	664	4 AAF22980	Aaf22980 Human pro
13	225	100.0	708	4 AAF22972	Aaf22972 Human pro
14	225	100.0	713	4 AAL10449	Aal10449 Human bre
15	225	100.0	713	4 AAF22974	Aaf22974 Human pro
16	225	100.0	741	4 AAF22991	Aaf22991 Human pro
17	225	100.0	766	4 AAF22986	Aaf22986 Human pro
18	225	100.0	771	4 AAL20493	Aal20493 Human bre
19	225	100.0	771	4 AAF22989	Aaf22989 Human pro
20	225	100.0	776	4 AAF22969	Aaf22969 Human pro
21	225	100.0	779	5 AAD06847	Aad06847 Human bre
22	225	100.0	814	4 AAF22967	Aaf22967 Human pro
23	225	100.0	833	4 AAF22987	Aaf22987 Human pro
24	225	100.0	843	11 ACN86163	Acn86163 Breast ca
25	225	100.0	898	11 ACN80647	Acn80647 Breast ca
26	225	100.0	899	11 ACN81796	Acn81796 Breast ca
27	225	100.0	904	4 AAF22965	Aaf22965 Human pro
28	225	100.0	1158	4 AAF22985	Aaf22985 Human pro
29	225	100.0	1337	4 AAF17979	Aaf17979 Human bre
30	225	100.0	1337	4 AAI67242	Aai67242 B726P spl
31	225	100.0	1337	4 AAS47409	Aas47409 Human cdn
32	225	100.0	1337	6 ABS64010	Abs64010 Human bre
33	225	100.0	1337	10 ABT33222	Abt33222 Human tum
34	225	100.0	1337	11 ADL93129	Adl93129 Human bre
35	225	100.0	1337	12 ADE44419	Ade44419 Human cdn
36	225	100.0	1448	5 ABV22558	Abv22558 Human pro
37	225	100.0	1448	5 ABV28376	Abv28376 Human pro
38	225	100.0	1448	5 ABV27350	Abv27350 Human pro
39	225	100.0	1478	11 ACN89282	Acn89282 Breast ca
40	225	100.0	1729	4 AAF17978	Aaf17978 Human bre
41	225	100.0	1729	4 AAI67221	Aai67221 B726P spl
42	225	100.0	1729	6 ABS64009	Abs64009 Human bre
43	225	100.0	1729	10 ABT33221	Abt33221 Human tum
44	225	100.0	1729	11 ADL93128	Adl93128 Human bre
45	225	100.0	1729	12 ADE44418	Ade44418 Human cdn

ALIGNMENTS

RESULT 1
AAAS9008
ID AAAS9008 standard; cDNA; 278 BP.
XX
XX
AC AAAS9008;
XX
XX 07-NOV-2000 (first entry)
XX
XX Nucleotide sequence of BS322 clone 4304443H1.
XX
XX BS322; breast tissue marker; breast disease; breast cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200043420-A1.
XX
XX 27-JUL-2000.
XX
XX 21-JAN-2000; 2000WO-US001452.
XX
XX 21-JAN-1999; 99US-00234716.
XX
XX (ABBO) ABBOTT LAB.
XX
XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
XX
XX Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
XX
XX Sroupe SD;
XX
XX WPI; 2000-499217/44.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
 PT detection and diagnosis of breast disease especially breast cancer.
 XX Claim 12; Fig 1A-E; 126pp; English.
 XX Clones AAAS9007-13 were used to produce the full length and consensus
 CC sequences encoding BS322 polypeptide. BS322 is a breast tissue marker.
 CC The BS322 polynucleotides and polypeptides are used to detect and
 CC diagnose breast disease, e.g. breast cancer. The BS322 polynucleotides
 CC are useful as a source of probes and primers, and the BS322 polypeptides
 CC are useful as antigens
 XX Sequence 278 BP; 104 A; 44 C; 62 G; 68 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.95e-25 Length: 278
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-489-079-26 (1-44) x AAA59008 (1-278)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 4 ATGCAAAAGTCTGTTCCTCAATTAAGCCTTGGAAATGAAGAAATGAACAACATTGAGAGCA 63
 QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 64 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTCTTGGGATTCT 123

QY 41 GluSerLeuCys 44
 Db 124 GAGAGTCTCTGT 135

RESULT 2

AAAL1595
 ID AAL1595 standard; cDNA; 445 BP.

XX
 AC AAL1595;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 4052.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US0000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

PS Claim 1; Page 740; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
 CC (AAU07344-AAU26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity
 XX Sequence 445 BP; 153 A; 83 C; 98 G; 109 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 7.25e-25 Length: 445
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAL1595 (1-445)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 39 ATGCAAAAGTCTGTTCCTCAATTAAGCCTTGGAAATGAAGAAATGAACAACATTGAGAGCA 98
 QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 99 GATGAGATACTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTCTTGGGATTCT 158

QY 41 GluSerLeuCys 44
 Db 159 GAGAGTCTCTGT 170

RESULT 3

AAAL16216

ID AAL16216 standard; cDNA; 533 BP.

XX
 AC AAL16216;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 8673.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US0000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 1570; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity
 XX
 SQ Sequence 533 BP; 188 A; 106 C; 117 G; 122 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9.15e-25 Length: 533
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAL16216 (1-533)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 |||||
 Db 67 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGATTGAAGAATGCAACAAACATTGAGAGCA 126
 |||||

QY 21 AspGluIleuProSerGluSerLysGlnLysAspTyGluGluSerSerTrpAspSer 40
 |||||
 Db 127 GATGAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 186
 |||||

QY 41 GluSerLeuCys 44
 |||||
 Db 187 GAGAGTCTCTGT 198
 |||||

RESULT 4
 AAL25059
 ID AAL25059 standard; cDNA; 562 BP.
 XX
 AC AAL25059;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 17516.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200151628-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 10-JAN-2001; 2001WO-US000798.
 XX
 PR 14-JAN-2000; 2000US-0176077P.
 PR 14-MAR-2000; 2000US-0189167P.
 PR 24-MAR-2000; 2000US-0192099P.
 PR 29-MAR-2000; 2000US-0193480P.
 PR 15-MAY-2000; 2000US-0205230P.
 PR 09-JUN-2000; 2000US-0211315P.
 PR 25-JUL-2000; 2000US-0220534P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lillie J, Xu Y, Wang Y, Steinmann K;
 XX
 DR WPI; 2001-451856/48.
 XX
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 PT
 XX Claim 1; Page 3237; 3695pp; English.
 PS
 XX The invention relates to human breast cancer expressed polynucleotides
 CC

CC The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity
 XX
 SQ Sequence 562 BP; 196 A; 105 C; 126 G; 135 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9.8e-25 Length: 562
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAL25059 (1-562)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 |||||
 Db 45 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGATTGAAGAATGCAACAAACATTGAGAGCA 104
 |||||

QY 21 AspGluIleuProSerGluSerLysGlnLysAspTyGluGluSerSerTrpAspSer 40
 |||||
 Db 105 GATGAGATACTCCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 164
 |||||

QY 41 GluSerLeuCys 44
 |||||
 Db 165 GAGAGTCTCTGT 176
 |||||

RESULT 5
 AAL20310/c
 ID AAL20310 standard; cDNA; 572 BP.
 XX
 AC AAL20310;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 12767.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200151628-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 10-JAN-2001; 2001WO-US000798.
 XX
 PR 14-JAN-2000; 2000US-0176077P.
 PR 14-MAR-2000; 2000US-0189167P.
 PR 24-MAR-2000; 2000US-0192099P.
 PR 29-MAR-2000; 2000US-0193480P.
 PR 15-MAY-2000; 2000US-0205230P.
 PR 09-JUN-2000; 2000US-0211315P.
 PR 25-JUL-2000; 2000US-0220534P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lillie J, Xu Y, Wang Y, Steinmann K;
 XX
 DR WPI; 2001-451856/48.
 XX
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 PT
 XX Claim 1; Page 2256; 3695pp; English.
 PS
 XX The invention relates to human breast cancer expressed polynucleotides
 CC

CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 572 BP; 137 A; 128 C; 109 G; 198 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1e-24 Length: 572
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAL20310 (1-572)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 528 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGAAATGAAGAATGAACAACATTGAGAGCA 469

QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 468 GATGAGTACTCCCATCAGATCCCAACAAAGGACTATGAAGAAGTTCTTGGGATTCT 409

QY 41 GluSerLeuCys 44
Db 408 GAGAGTCTCTGT 397

RESULT 6
AAL20029
ID AAL20029 standard; cDNA; 578 BP.
XX
AC AAL20029;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 12486.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US0000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
PS Claim 1; Page 2207-2208; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 578 BP; 201 A; 112 C; 124 G; 141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.02e-24 Length: 578
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAL20029 (1-578)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 57 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGAAATGAAGAATGAACAACATTGAGAGCA 116

QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 117 GATGAGTACTCCCATCAGATCCCAACAAAGGACTATGAAGAAGTTCTTGGGATTCT 176

QY 41 GluSerLeuCys 44
Db 177 GAGAGCTCTGT 188

RESULT 7
AAH55638/c
ID AAH55638 standard; DNA; 580 BP.
XX
AC AAH55638;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human breast tumour protein clone 48970 DNA sequence.
XX
KW Cytostatic; vaccine; human; breast tumour protein; breast cancer;
KW gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200140269-A2.
XX
PD 07-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US032520.
XX
PR 30-NOV-1999; 98US-00451651.
PR 22-FEB-2000; 2000US-00510662.
PR 10-MAR-2000; 2000US-00523586.
PR 07-APR-2000; 2000US-00545068.
PR 15-MAY-2000; 2000US-00571025.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang A;
XX
DR WPI; 2001-356154/37.
XX
PT Breast tumor polypeptides and the nucleic acids that encode them, useful
PT for the prevention, diagnosis and treatment of breast cancer.
XX
PS Claim 24; Page 172; 221pp; English.
XX
CC The present sequence is a human breast tumour protein coding sequence.
CC This sequence may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the breast tumour

CC protein e.g. breast cancer. For example, this sequence may be used to
 CC treat disorders associated with decreased expression by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC breast tumour protein by expressing inactive proteins or to supplement
 CC the patient's own production of the breast tumour protein. Additionally,
 CC the present sequence may be used to produce the breast tumour protein, by
 CC inserting the nucleic acids into a host cell and culturing the cell to
 CC express the protein. The present sequence and its complementary sequences
 CC may also be used as DNA probes in diagnostic assays to detect, and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy
 XX
 SQ Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;

Alignment Scores:
 Pred. No.: 1,02e-24 Length: 580
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AH55638 (1-580)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 DB 523 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGAAATGAAGATGCAACAACATTGAGAGCA 464
 QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrcGluLysSerSerTrpAspSer 40
 DB 463 GATGAGATATCTCCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 404
 QY 41 GluSerLeuCys 44
 DB 403 GAGAGTCTCTGT 392

RESULT 8

ID ADN40408/C
 ID ADN40408 standard; cDNA; 580 BP.

XX ADN40408;

XX 12-AUG-2004 (first entry)

XX Human breast cancer associated cDNA sequence #160.

XX Human; breast cancer; T cell; tumour protein; antigen presenting cell;
 KW immune response; CD4+; CD8+; cytostatic; gene therapy; ss.

XX Homo sapiens.

XX US2004101899-A1.

XX 27-MAY-2004.

XX 13-NOV-2003; 2003US-00714389.

XX 30-NOV-1999; 99US-00451651.

XX 22-FEB-2000; 2000US-00510662.

XX 10-MAR-2000; 2000US-00523586.

XX 07-APR-2000; 2000US-00545068.

XX 15-MAY-2000; 2000US-00571025.

XX 06-FEB-2001; 2001US-00778320.

XX (CORI-) CORIXA CORP.

XX Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;

PI McNeill PD;

XX WPI; 2004-399689/37.

XX New polynucleotides, useful for treating and diagnosing cancer,
 PT particularly breast cancer by stimulating immune response in a patient

PT and inhibiting the development of cancer.

XX Claim 1; SEQ ID NO 163; 151pp; English.

XX The present invention relates to polynucleotide and polypeptide sequences
 CC associated with breast cancer. Also disclosed are expression vectors
 CC comprising the polynucleotide sequences of the invention operably linked
 CC to an expression control sequence, host cells comprising the vector,
 CC antibodies (or antigen binding fragments of antibodies) specifically
 CC binding the polypeptides of the invention, fusion proteins comprising at
 CC least one of the polypeptides, stimulating and/or expanding T cells
 CC specific for a tumour protein. The polynucleotide sequences, polypeptide
 CC sequences, and antigen presenting cells can be administered
 CC therapeutically/prophylactically to induce an immune response. They can
 CC be included with a physiological carrier/immunostimulant in compositions
 CC such as vaccines, particularly to treat or prevent cancers such as breast
 CC cancer. They can also be used to inhibit the development of cancer by
 CC incubating one or more of them with CD4+ and/or CD8+ T cells isolated
 CC from a patient, such that the T cells proliferate, and administering the
 CC proliferated T cells to the patient. The polynucleotide sequences are
 CC useful for detecting cancer in a patient, producing fusion proteins,
 CC producing T cell populations and antigen presenting cells. The present
 CC sequence represents a polynucleotide sequence of the invention.

XX Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 1,02e-24 Length: 580
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-489-079-26 (1-44) x ADN40408 (1-580)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 DB 523 ATGCAAAAGTCTGTTCCAAATAAAGCTTGGAAATGAAGATGCAACAACATTGAGAGCA 464

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrcGluLysSerSerTrpAspSer 40
 DB 463 GATGAGATATCTCCCATCAGAAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCT 404

QY 41 GluSerLeuCys 44

DB 403 GAGAGTCTCTGT 392

RESULT 9

ID ACN92774

XX ACN92774 standard; DNA; 600 BP.

XX ACN92774;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 13924.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.
XX Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 13924; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with a
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 600 BP; 205 A; 108 C; 130 G; 136 T; 0 U; 21 Other;

Alignment Scores:
Pred. No.: 1,07e-24 Length: 600
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-489-079-26 (1-44) x ACN92774 (1-600)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 436 ATGCAAAAGTCGTGTTCCAAATTAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCA 495
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 496 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTTCTTGGGATTCT 555

QY 41 GluSerLeuCys 44
Db 556 GAGAGTCCTCTGT 567
RESULT 10
ACN87508
ID ACN87508 standard; DNA; 635 BP.
XX
XX ACN87508;
XX
XX 02-DEC-2004 (first entry)
XX
XX Breast cancer related marker, seq id 8658.
XX
XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX Homo sapiens.
XX
XX US2003099974-A1.
XX
XX 29-MAY-2003.
XX
XX 18-JUL-2002; 2002US-00198846.
XX
XX 18-JUL-2001; 2001US-0306220P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for
PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
XX Disclosure; SEQ ID NO 8658; 36pp; English.
XX
XX The invention relates to an isolated polypeptide (I) associated with a
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 635 BP; 213 A; 128 C; 154 G; 135 T; 0 U; 5 Other;

Alignment Scores:
Pred. No.: 1.15e-24 Length: 635
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-489-079-26 (1-44) x ACN87508 (1-635)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 378 ATGCAAAAGTCGTGTTCCAAATTAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCA 437
QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 438 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAGTTCTTGGGATTCT 497
QY 41 GluSerLeuCys 44
Db 498 GAGAGTCCTCTGT 509

RESULT 11
AAF22981
ID AAF22981 standard; cDNA; 636 BP.
XX
XX AAF22981;
XX
XX 26-MAR-2001 (first entry)
XX
XX Human prostate cancer associated antigen nucleotide sequence SEQ ID:560.
XX
XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine; ss.
XX
XX Homo sapiens.
XX
XX WO200073801-A2.
XX
XX 07-DEC-2000.
XX
XX 26-MAY-2000; 2000WO-US014749.
XX
XX 28-MAY-1999; 99US-0136526P.
XX
XX 10-SEP-1999; 99US-0153454P.
XX
XX (LUDW-) LUDWIG INST CANCER RES.
XX
XX Obata Y;
XX
XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Claim 50; Page 456; 799pp; English.
 XX
 CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 SQ Sequence 636 BP; 217 A; 120 C; 138 G; 157 T; 0 U; 4 Other;

Alignment Scores:
 Pred. No.: 1.15e-24 Length: 636
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAF22981 (1-636)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 DB 234 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGAATTGAAGAAATGAACAACATTGAGAGCA 293
 QY 21 AspGluIleuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 DB 294 GATGAGATACTCCCATCAGAATCCAAACAAAGAGGACTATGAGAAAGTTCTTGGGATTCT 353
 QY 41 GluSerLeuCys 44
 DB 354 GAGAGTCTCTGT 365

RESULT 12
 AAF22980
 ID AAF22980 standard; cDNA; 664 BP.
 XX
 AC AAF22980;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen nucleotide sequence SEQ ID:559.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 DR WPI; 2001-025274/03.
 XX

PT Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 PS Claim 50; Page 455; 799pp; English.
 XX
 CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 SQ Sequence 664 BP; 236 A; 115 C; 144 G; 166 T; 0 U; 3 Other;

Alignment Scores:
 Pred. No.: 1.22e-24 Length: 664
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAF22980 (1-664)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 DB 24 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGAATTGAAGAAATGAACAACATTGAGAGCA 83
 QY 21 AspGluIleuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 DB 84 GATGAGATACTCCCATCAGAATCCAAACAAAGAGGACTATGAGAAAGTTCTTGGGATTCT 143
 QY 41 GluSerLeuCys 44
 DB 144 GAGAGTCTCTGT 155

RESULT 13
 AAF22972
 ID AAF22972 standard; cDNA; 708 BP.
 XX
 AC AAF22972;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen nucleotide sequence SEQ ID:551.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 DR WPI; 2001-025274/03.
 XX
 PT Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.

XX
XX
PS Claim 50; Page 453; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer

XX
SQ Sequence 708 BP; 246 A; 134 C; 154 G; 172 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 1.32e-24 Length: 708
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAF22972 (1-708)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 271 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGAATTGAAGATGAACAACATTGAGAGCA 330

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 331 GATGAGATACCTCCATCAGATCCAAACAAAGAGGACTATGAAGAAGTCTTGGGATTCT 390

QY 41 GluSerLeuCys 44
Db 391 GAGAGTCTCTGT 402

RESULT 14
AAL10449
ID AAL10449 standard; cDNA; 713 BP.
XX
XX AAL10449;
AC AAL10449;
XX
XX 07-DEC-2001 (first entry)
DT Human breast cancer expressed polynucleotide 2906.
DE Human; breast cancer; cell marker; cytostatic; ss.
KW Homo sapiens.
XX
XX WO200151628-A2.
PN
XX 19-JUL-2001.
PD
XX 10-JAN-2001; 2001WO-US000798.
PF
XX 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Lillie J, Xu Y, Wang Y, Steinmann K;
PI
XX

DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX PS Claim 1; Page 542; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

XX
SQ Sequence 713 BP; 224 A; 148 C; 168 G; 167 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 1.33e-24 Length: 713
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAL10449 (1-713)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 100 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGAATTGAAGATGAACAACATTGAGAGCA 159

QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 160 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAGTCTTGGGATTCT 219

QY 41 GluSerLeuCys 44
Db 220 GAGAGCTCTCTGT 231

RESULT 15
AAF22974
ID AAF22974 standard; cDNA; 713 BP.
XX
XX AAF22974;
AC AAF22974;
XX
XX 26-MAR-2001 (first entry)
DT Human prostate cancer associated antigen nucleotide sequence SEQ ID:553.
DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine; ss.
XX
XX Homo sapiens.
OS
XX WO200073801-A2.
PN
XX 07-DEC-2000.
PD
XX 26-MAY-2000; 2000WO-US014749.
PF
XX 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX (LUDW-) LUDWIG INST CANCER RES.
PA
XX Obata Y;
PI
XX WPI; 2001-025274/03.
DR
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.

PT cancer.

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Claim 50; Page 453-454; 799pp; English.

AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

represent nucleotide sequences encoding human breast, gastric and

prostate cancer associated antigen precursors (CAAP) respectively.

AAF63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

represent human breast, gastric and prostate CAAP protein sequence

respectively. CAAPs have cytostatic activity and can be used in the

production of cancer vaccines. The human CAAP proteins, peptides, nucleic

acids or anti-CAAP antibodies are useful for diagnosing and treating a

condition characterised by expression of an abnormal amount of a protein,

e.g. cancer

Sequence 713 BP; 250 A; 133 C; 153 G; 172 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.: 1.33e-24 Length: 713

Score: 225.00 Matches: 44

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-09-489-079-26 (1-44) x AAF22974 (1-713)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20

DB 319 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGAATTCGAAGAATGCAACAAACATTGAGAGCA 378

QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40

DB 379 GATGAGATATCTCCATCAGATCCAAACAAAGGACTATGAGCAAGATCTTGGGATTCT 438

QY 41 GluSerLeuCys 44

DB 439 GAGAGTCTCTGT 450

Search completed: August 1, 2005, 23:18:11

Job time : 85.4834 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:44:18 ; Search time 1673.51 Seconds
(without alignments)
1000.786 Million cell updates/sec

Title: US-09-489-079-26
Perfect score: 225
Sequence: 1 MOKSVNKALEKNEQTLRA.....PSEKQDYBESSWDSLSIC 44

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09489079/runat_01082005_141250_13309/app_query.fasta_1.1635
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CNC 1.1 10973 @runat_01082005_141250_13309 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gse1.*
9: gb_gse2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	225	100.0	581	5	BP314867
2	225	100.0	582	5	BP313026
3	225	100.0	582	5	BP313235
4	225	100.0	582	5	BP313436
5	225	100.0	582	5	BP313704
6	225	100.0	582	5	BP314260
7	225	100.0	582	5	BP315806
8	225	100.0	583	5	BP312890
9	225	100.0	583	5	BP313822

10	225	100.0	584	5	BP314652	BP314652	BP314652
11	225	100.0	598	5	BP312910	BP312910	BP312910
12	212	94.2	443	2	BF746270	BF746270	RC1-BF7025
13	212	94.2	443	2	BF746340	BF746340	RC1-BF7025
14	210	93.3	583	5	BP313800	BP313800	BP313800
15	199	88.4	582	1	BP315089	BP315089	BP315089
c	16	190	84.4	490	1	AI951118	wx63905.x
17	180	80.0	3443	3	BC028407	BC028407	Homo sapi
18	175	77.8	574	2	BF328582	BF328582	RC5-BN019
c	19	122	54.2	563	8	AQ546028	CITBI-E1-
20	97	43.1	165	9	HSMC18H11	H88328	H. sapiens D
21	97	43.1	514	5	BP312630	BP312630	BP312630
c	22	93	41.3	705	8	AQ423531	CITBI-E1-
23	92.5	41.1	344	8	AZ576125	AQ2576125	AST-T32E0
24	88	39.1	389	7	CR735535	CR735535	CR735535
25	88	39.1	596	5	BQ429618	BQ429618	AGENCOURT
26	85.5	38.0	155	2	BF171278	BF171278	PCL2374 M
c	27	78.5	34.9	847	1	AU132223	AU132223
28	78.5	34.9	5193	3	CR627244	CR627244	Homo sapi
29	74	32.9	431	2	BE062233	BE062233	RC1-BT025
30	67.5	30.0	535	5	BH811432	BH811432	UL84TC04
31	67.5	30.0	550	7	CK102592	CK102592	G084P07.5
32	67.5	30.0	553	4	BI128959	BI128959	G084P07Y
33	67	29.8	655	7	CN787741	CN787741	4121950 B
34	66	29.3	578	5	BP349185	BP349185	BP349185
35	66	29.3	732	7	CN841887	CN841887	AGENCOURT
c	36	65.5	29.1	963	2	BF179023	BF179023
37	65.5	29.1	1069	9	CL109041	CL109041	ISB1-S0N1
c	38	65.5	29.1	1075	9	CL108568	ISB1-S0N1
39	65	28.9	349	9	CE315218	CE315218	tigr-988-
c	40	65	28.9	530	8	AZ110334	RPCI-23-4
c	41	65	28.9	539	1	AL917950	AL917950
42	65	28.9	771	7	CK690602	CK690602	ZF101-P00
43	65	28.9	3527	3	HSM801761	ALU136793	Homo sapi
44	64.5	28.7	828	9	CC485421	CC485421	CH240_314
c	45	64.5	1039	9	CL018306	CL018306	CH216-3G5

ALIGNMENTS

RESULT 1
BP314867
LOCUS
DEFINITION
BP314867 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
CDNA clone OFR07276, mRNA sequence.
581 bp mRNA linear EST 17-SEP-2004

ACCESSION
BP314867
VERSION
EST.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 581)
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
JOURNAL
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
Location/Qualifiers
source 1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR07276"
/issue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 2,29e-21 Length: 581
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP314867 (1-581)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 43 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGATTGAAGATGAACAACATTGAGAGCA 102
 Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 103 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 162
 Qy 41 GluSerLeuCys 44
 Db 163 GAGAGTCTCTGT 174

RESULT 2
 BP313026
 LOCUS BP313026 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION BP313026 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 ACCESSION BP313026
 VERSION BP313026.1 GI:52242001
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES
 source
 1..582
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OPR01939"
 /tissue_type="mammary gland"
 /cell_line="OCUB-F"
 /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 2,3e-21 Length: 582
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP313026 (1-582)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 346 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGATTGAAGATGAACAACATTGAGAGCA 405

Qy

Db 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 406 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 465
 Qy 41 GluSerLeuCys 44
 Db 466 GAGAGTCTCTGT 477

RESULT 3

BP313235

LOCUS BP313235 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION BP313235 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 ACCESSION BP313235
 VERSION BP313235.1 GI:52242210
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES

source
 1..582
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OPR02569"
 /tissue_type="mammary gland"
 /cell_line="OCUB-F"
 /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 2,3e-21 Length: 582
 Score: 225.00 Matches: 44
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP313235 (1-582)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
 Db 244 ATGCAAAAGTCTGTTCCAAATAAAGCCTTGGATTGAAGATGAACAACATTGAGAGCA 303

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
 Db 304 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 363

Qy

Db 41 GluSerLeuCys 44

364 GAGAGTCTCTGT 375

RESULT 4

BP313436

LOCUS BP313436 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION BP313436 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 ACCESSION BP313436
 VERSION BP313436.1 GI:52242411
 KEYWORDS EST.

```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 582)
AUTHORS    Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
            Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES   source
            1..582
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="OFRO3209"
            /tissue_type="mammary gland"
            /cell_line="OCUB-F"
            /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
            /note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.:      2,3e-21      Length:      582
Score:          225.00      Matches:     44
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00%     Indels:      0
DB:             5          Gaps:         0

US-09-489-079-26 (1-44) x BP313704 (1-582)

QY      1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
        |||||
DB      371 ATGCAAAAGCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 98
        |||||

QY      21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
        |||||
DB      431 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTCT 490
        |||||

QY      41 GluSerLeuCys 44
        |||||
DB      491 GAGAGTCTCTGT 502
        |||||

RESULT 5
BP313704
LOCUS     BP313704      582 bp      mRNA      linear      EST 17-SEP-2004
DEFINITION
            CDNA clone OFRO3974, mammary gland OCUB-F Homo sapiens
ACCESSION
            BP313704
VERSION    BP313704.1 GI:52242679
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 582)
AUTHORS    Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
            Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES   source
            1..582
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="OFRO3209"
            /tissue_type="mammary gland"
            /cell_line="OCUB-F"
            /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
            /note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.:      2,3e-21      Length:      582
Score:          225.00      Matches:     44
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00%     Indels:      0
DB:             5          Gaps:         0

US-09-489-079-26 (1-44) x BP313436 (1-582)

QY      1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
        |||||
DB      371 ATGCAAAAGCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 430
        |||||

QY      21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
        |||||
DB      431 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTCT 490
        |||||

QY      41 GluSerLeuCys 44
        |||||
DB      491 GAGAGTCTCTGT 502
        |||||

RESULT 5
BP313704
LOCUS     BP313704      582 bp      mRNA      linear      EST 17-SEP-2004
DEFINITION
            CDNA clone OFRO3974, mammary gland OCUB-F Homo sapiens
ACCESSION
            BP313704
VERSION    BP313704.1 GI:52242679
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 582)
AUTHORS    Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
            Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES   source
            1..582
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="OFRO3209"
            /tissue_type="mammary gland"
            /cell_line="OCUB-F"
            /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
            /note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.:      2,3e-21      Length:      582
Score:          225.00      Matches:     44
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00%     Indels:      0
DB:             5          Gaps:         0

US-09-489-079-26 (1-44) x BP313704 (1-582)

QY      1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
        |||||
DB      39 ATGCAAAAGCTGTTCCAAATAAAGCTTGGATTGAAGATGAACAAACATTGAGAGCA 98
        |||||

QY      21 AspGluLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
        |||||
DB      99 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTTGGGATTCT 158
        |||||

QY      41 GluSerLeuCys 44
        |||||
DB      159 GAGAGTCTCTGT 170
        |||||

RESULT 6
BP314260
LOCUS     BP314260      582 bp      mRNA      linear      EST 17-SEP-2004
DEFINITION
            CDNA clone OFRO5703, mRNA sequence.
ACCESSION
            BP314260
VERSION    BP314260.1 GI:52243235
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 582)
AUTHORS    Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
            Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE      Sequence comparison of human and mouse genes reveals a homologous
            block structure in the promoter regions
JOURNAL    Genome Res. 14 (9), 1711-1718 (2004)
COMMENT    Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES   source
            1..582
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="OFRO5703"
            /tissue_type="mammary gland"
            /cell_line="OCUB-F"
            /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
            /note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.:      2,3e-21      Length:      582
Score:          225.00      Matches:     44
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```

Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP314260 (1-582)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 246 ATGCAAAAGTCTGTTCCAAATAGGCTTGGAAATGAAGAATGAACAACATTGAGAGCA 305

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTriaAspSer 40
Db 306 GATGAGATATCTCCATCAGATCAATCAACAAAGGACTATGAGAAGTCTTGGGATCT 365

Qy 41 GluSerLeuCys 44
Db 366 GAGAGTCTCTGT 377

RESULT 7
BP315806 582 bp mRNA linear EST 17-SEP-2004
LOCUS
DEFINITION
BP315806 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION
VERSION OFR09418, mRNA sequence.
KEYWORDS
SOURCE
BP315806.1 GI:52244781
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 582)
AUTHORS
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL
Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
Location/Qualifiers
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR09418"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.: 2,3e-21 Length: 582
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP315806 (1-582)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 246 ATGCAAAAGTCTGTTCCAAATAGGCTTGGAAATGAAGAATGAACAACATTGAGAGCA 305

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTriaAspSer 40
Db 306 GATGAGATATCTCCATCAGATCAATCAACAAAGGACTATGAGAAGTCTTGGGATCT 365

Qy 41 GluSerLeuCys 44
Db 366 GAGAGTCTCTGT 377

RESULT 8
BP312890 583 bp mRNA linear EST 17-SEP-2004
LOCUS
DEFINITION
BP312890 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION
VERSION OFR01356, mRNA sequence.
KEYWORDS
SOURCE
BP312890.1 GI:52241865
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 583)
AUTHORS
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL
Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
Location/Qualifiers
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR01356"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.: 2,3e-21 Length: 583
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP312890 (1-583)

Qy 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 260 ATGCAAAAGTCTGTTCCAAATAGGCTTGGAAATGAAGAATGAACAACATTGAGAGCA 319

Qy 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTriaAspSer 40
Db 320 GATGAGATATCTCCATCAGATCAATCAACAAAGGACTATGAGAAGTCTTGGGATCT 379

Qy 41 GluSerLeuCys 44
Db 380 GAGAGTCTCTGT 391

RESULT 9
BP313822 583 bp mRNA linear EST 17-SEP-2004
LOCUS
DEFINITION
BP313822 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION
VERSION OFR04566, mRNA sequence.
KEYWORDS
SOURCE
BP313822.1 GI:52242797
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 583)
AUTHORS
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.

```

TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OF04566"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 2,3e-21 Length: 583
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP313822 (1-583)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlnThrIleuArgAla 20
Db 215 ATGCAAAAGTCTGTTCCAAATAAGCCTTGGAATTAAGCAATGAACAAACGTTGAGAGCA 274

QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluLysSerSerTrpAspSer 40
Db 275 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 334

QY 41 GluSerLeuCys 44
Db 335 GAGAGTCTCTGT 346

RESULT 10
LOCUS BP314652
DEFINITION BP314652 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION BP314652
VERSION BP314652.1 GI:52243627
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 584)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .584
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OF06752"
/tissue_type="mammary gland"
/cell_line="OCUB-F"

/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 2,31e-21 Length: 584
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP314652 (1-584)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlnThrIleuArgAla 20
Db 260 ATGCAAAAGTCTGTTCCAAATAAGCCTTGGAATTAAGCAATGAACAAACATTTGAGAGCA 319

QY 21 AspGluIleLeuProSerGluSerLysGlnLysAspTyrGluLysSerSerTrpAspSer 40
Db 320 GATGAGATACCTCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 379

QY 41 GluSerLeuCys 44
Db 380 GAGAGTCTCTGT 391

RESULT 11
LOCUS BP312910
DEFINITION BP312910 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION BP312910
VERSION BP312910.1 GI:52241885
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 598)
AUTHORS Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1. .598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OF01441"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 2,37e-21 Length: 598
Score: 225.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-26 (1-44) x BP312910 (1-598)

QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGlnThrIleuArgAla 20
|||||

```

Db      261 ATGCAAAAGTCGTGTTCCAAATAAAGCCTTTGGAAATTGAAGATGAACAACAATTGAGAGCA 320
QY      21 AspGlu1LeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
      |||||
Db      321 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 380
QY      41 GluSerLeuCys 44
      |||||
Db      381 GAGAGTCTCTGT 392

RESULT 12
LOCUS   BF746270                      443 bp      mRNA      linear      EST 10-JAN-2001
DEFINITION RC1-BT0254-071100-119-a03 BT0254 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF746270
VERSION   BF746270.1 GI:12072946
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          10737800
          PUBMED
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC1&t2=RC1-BT0254-
          071100-119-a03&t3=2000-11-07&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 9
          High quality sequence stop: 439.
          Location/Qualifiers
FEATURES             source
          1..443
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="BT0254"
             /note="Organ: breast; Vector: puc18; Site 1: Smal; Site 2:
             Smal; A mini-library was made by cloning products derived
             from ORESTES PCR (U.S. Letters Patent application No.
             196,716 - Ludwig Institute for Cancer Research) profiles
             into the pUC 18 vector. Reverse transcription of tissue
             mRNA and cDNA amplification were performed under low
             stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.:      1,12e-19      Length:      443
Score:          212.00      Matches:      43
Percent Similarity: 97.73%      Conservative: 0
Best Local Similarity: 97.73%      Mismatches: 1
Query Match:    94.22%      Indels:      0
Gaps:           2

US-09-489-079-26 (1-44) x BF746270 (1-443)

```

```

QY      1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
      |||||
Db      178 ATGCAAAAGTCGTGTTCCAAATAAAGCCTTTGGAAATTGAAGATGAACAACAATTGAGAGCA 237
QY      21 AspGlu1LeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
      |||||
Db      238 GATGAGATACCTCCCATCAGATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCT 297
QY      41 GluSerLeuCys 44
      |||||
Db      298 GAGAGTCTCTGT 309

RESULT 13
LOCUS   BF746340                      443 bp      mRNA      linear      EST 10-JAN-2001
DEFINITION RC1-BT0254-081100-119-a03 BT0254 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF746340
VERSION   BF746340.1 GI:12073016
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          10737800
          PUBMED
COMMENT   Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC1&t2=RC1-BT0254-
          081100-119-a03&t3=2000-11-08&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 9
          High quality sequence stop: 439.
          Location/Qualifiers
FEATURES             source
          1..443
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="BT0254"
             /note="Organ: breast; Vector: puc18; Site 1: Smal; Site 2:
             Smal; A mini-library was made by cloning products derived
             from ORESTES PCR (U.S. Letters Patent application No.
             196,716 - Ludwig Institute for Cancer Research) profiles
             into the pUC 18 vector. Reverse transcription of tissue
             mRNA and cDNA amplification were performed under low
             stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.:      1,12e-19      Length:      443
Score:          212.00      Matches:      43
Percent Similarity: 97.73%      Conservative: 0
Best Local Similarity: 97.73%      Mismatches: 1
Query Match:    94.22%      Indels:      0

```

```
DB: 2 Gaps: 0
US-09-489-079-26 (1-44) x BF746340 (1-443)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 178 ATGCAAAAGTCTGTTCCTCAATAAAGCCTTGGAATTGAAGATGAACAAACATTGAGAGCA 237
QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 238 GATGAGATACTCCCATCAGATCCAAATCCAAAGGACTATGAGAAAGTTCTGGGATTTCT 297
QY 41 GluSerLeuCys 44
Db 298 GAGAGTCTCTGT 309

RESULT 14
BP313800
LOCUS
DEFINITION
BP313800 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION
BP313800
VERSION
BP313800.1 GI:52242775
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 583)
AUTHORS
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL
Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR04504"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"
ORIGIN
Alignment Scores:
Pred. No.: 2,99e-19 Length: 583
Score: 210.00 Matches: 43
Percent Similarity: 97.73% Conservative: 0
Best Local Similarity: 97.73% Mismatches: 1
Query Match: 93.33% Indels: 0
DB: 5 Gaps: 0
US-09-489-079-26 (1-44) x BP313800 (1-583)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 242 ATGCAAAAGTCTGTTCCTCAATAAAGCCTTGGAATTGAAGATGAACAAACATTGAGAGCA 301
QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 302 GATGAGATACTCCCATCAGATCCAAATCCAAAGGACTATGAGAAAGTTCTTGAGATTTCT 361
QY 41 GluSerLeuCys 44
Db 362 GAGAGTCTCTGT 373

RESULT 15
BP315089
LOCUS
DEFINITION
BP315089 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION
BP315089
VERSION
BP315089.1 GI:52244064
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 582)
AUTHORS
Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL
Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR07766"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"
ORIGIN
Alignment Scores:
Pred. No.: 1,06e-17 Length: 582
Score: 199.00 Matches: 40
Percent Similarity: 93.18% Conservative: 1
Best Local Similarity: 90.91% Mismatches: 3
Query Match: 88.44% Indels: 0
DB: 5 Gaps: 0
US-09-489-079-26 (1-44) x BP315089 (1-582)
QY 1 MetGlnLysSerValProAsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAla 20
Db 304 ATGCAAAAGTCTGTTCCTCAATAAAGCCTTGGAATTGAAGATGAACAAACATTGAGAGCA 363
QY 21 AspGluLeuLeuProSerGluSerLysGlnLysAspTyrGluGluSerSerTrpAspSer 40
Db 364 GATGAGATACTCCCATCAGATCCAAATCCAAAGGACTATGAGAAAGTTCTTGATTTCT 423
QY 41 GluSerLeuCys 44
Db 424 GAGAGTCTCTGT 435

Search completed: August 2, 2005, 11:46:11
Job time : 1677.51 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:23:28 ; Search time 70.372 Seconds
(without alignments)
3196.585 Million cell updates/sec

Title: US-09-489-079-27
Perfect score: 195
Sequence: 1 NKAELKNEQTLRADEILPESKQDYBESSWDSLSLC 38

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USFTO.spool_p/US0489079/runat_01082005_141249_13290/app_query.fasta_1.1635
-DB=N Geneseq 16Dec04 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0489079 @CGN 1.1 1418 @runat_01082005_141249_13290 -NCFU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	Geneseqn16Dec04:*
2:	Geneseqn1980s:*
3:	Geneseqn1998s:*
4:	Geneseqn2000s:*
5:	Geneseqn2001as:*
6:	Geneseqn2002as:*
7:	Geneseqn2002bs:*
8:	Geneseqn2003as:*
9:	Geneseqn2003bs:*
10:	Geneseqn2003cs:*
11:	Geneseqn2003ds:*
12:	Geneseqn2004as:*
13:	Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	195	100.0	278	3	Aaa59008 Nucleotid
2	195	100.0	445	4	Aall1595 Human bre
3	195	100.0	466	5	Abv13085 Human pro
4	195	100.0	473	5	Abv43067 Human pro
5	195	100.0	473	5	Abv34206 Human pro

6	195	100.0	508	5	ABV03916 Human pro
7	195	100.0	533	4	AAL16216 Human bre
8	195	100.0	562	4	AAL25059 Human bre
9	195	100.0	572	4	AAL20310 Human bre
10	195	100.0	578	4	AAL20029 Human bre
11	195	100.0	580	4	Aah55638 Human bre
12	195	100.0	580	12	Adn40408 Human bre
13	195	100.0	600	11	Acn92774 Breast ca
14	195	100.0	635	11	Acn87508 Breast ca
15	195	100.0	636	4	Aaf22981 Human pro
16	195	100.0	664	4	Aaf22980 Human pro
17	195	100.0	708	4	Aaf22972 Human pro
18	195	100.0	713	4	Aal10449 Human bre
19	195	100.0	713	4	Aaf22974 Human pro
20	195	100.0	741	4	Aaf22991 Human pro
21	195	100.0	766	4	Aaf22986 Human pro
22	195	100.0	771	4	AAL20493 Human bre
23	195	100.0	771	4	Aaf22989 Human pro
24	195	100.0	776	4	Aaf22969 Human pro
25	195	100.0	779	5	AAD06847 Human bre
26	195	100.0	814	4	Aaf22967 Human pro
27	195	100.0	833	4	Aaf22987 Human pro
28	195	100.0	843	11	ACN86163 Breast ca
29	195	100.0	898	11	ACN80647 Breast ca
30	195	100.0	899	11	ACN81796 Breast ca
31	195	100.0	904	4	Aaf22965 Human pro
32	195	100.0	1158	4	Aaf22985 Human pro
33	195	100.0	1337	4	Aaf17979 Human bre
34	195	100.0	1337	4	Aai67222 B726P spl
35	195	100.0	1337	4	Aas47409 Human cdn
36	195	100.0	1337	6	ABs64010 Human bre
37	195	100.0	1337	10	ABT33222 Human tum
38	195	100.0	1337	11	ADL93129 Human bre
39	195	100.0	1337	12	ADE44419 Human cdn
40	195	100.0	1448	5	ABV22558 Human pro
41	195	100.0	1448	5	ABV28376 Human pro
42	195	100.0	1448	5	ABV27350 Human pro
43	195	100.0	1478	11	ACN89282 Breast ca
44	195	100.0	1729	4	Aaf17978 Human bre
45	195	100.0	1729	4	Aai67221 B726P spl

ALIGNMENTS

RESULT 1
AAA59008
ID AAA59008 standard; cdna; 278 BP.
XX
AC AAA59008;
XX
DT 07-NOV-2000 (first entry)
XX
DE Nucleotide sequence of BS322 clone 4304443H1.
XX
KW BS322; breast tissue marker; breast disease; breast cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200043420-A1.
XX
PD 27-JUL-2000.
XX
PF 21-JAN-2000; 2000WO-US001452.
XX
PR 21-JAN-1999; 99US-00234716.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
PI Stroupe SD;
XX
DR WPI; 2000-499217/44.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
PT detection and diagnosis of breast disease especially breast cancer.
XX
XX Claim 12; Fig 1A-E; 126pp; English.

XX Clones AAA59007-13 were used to produce the full length and consensus
CC sequences encoding BS322 polypeptide. BS322 is a breast tissue marker.
CC The BS322 polynucleotides and polypeptides are used to detect and
CC diagnose breast disease, e.g. breast cancer. The BS322 polynucleotides
CC are useful as a source of probes and primers, and the BS322 polypeptides
CC are useful as antigens

SQ Sequence 278 BP; 104 A; 44 C; 62 G; 68 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2e-21 Length: 278
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-489-079-27 (1-38) x AAA59008 (1-278)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
Db 22 AATAAGCCTTGGAAATGGAATGCAACAAATTCGAGACGAGATGAGATCTCCATCA 81

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 82 GAATCCAAACAAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 135

RESULT 2

AAL11595
ID AAL11595 standard; cDNA; 445 BP.

XX AAL11595;

XX 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 4052.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

OS WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 740; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is

CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

SQ Sequence 445 BP; 153 A; 83 C; 98 G; 109 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3.69e-21 Length: 445
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAL11595 (1-445)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
Db 57 AATAAGCCTTGGAAATGGAATGCAACAAATTCGAGACGAGATGAGATCTCCATCA 116

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 117 GAATCCAAACAAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 170

RESULT 3

ABV13085

ID ABV13085 standard; cDNA; 466 BP.

XX ABV13085;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 13076.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 2161; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 466 BP; 171 A; 87 C; 104 G; 104 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3-92e-21 Length: 466
 Score: 195.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x ABV13085 (1-466)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
 DB 21 AATAAGCCCTTGAATTCGAATGAATGAACAAACATTGAGAGCAGATGATATCCCATCA 80
 QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
 DB 81 GAATCCAAACAAAGGACTATGAAGAAAGTCTCTGGATTCTGAGAGTCTCTGT 134

RESULT 4
 ABV43067
 ID ABV43067 standard; cDNA; 473 BP.

XX AC ABV43067;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 43058.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 8599-8600; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 473 BP; 162 A; 91 C; 115 G; 105 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4e-21 Length: 473
 Score: 195.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x ABV43067 (1-473)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
 DB 63 AATAAGCCCTTGAATTCGAATGAACAAACATTGAGAGCAGATGATATCCCATCA 122

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38

DB 123 GAATCCAAACAAAGGACTATGAAGAAAGTCTCTGGATTCTGAGAGTCTCTGT 176

RESULT 5
 ABV43206

ID ABV43206 standard; cDNA; 473 BP.

XX AC ABV43206;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 34197.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 7198; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 473 BP; 162 A; 91 C; 115 G; 105 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4e-21 Length: 473
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x ABV34206 (1-473)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
Db 63 AATAAGCCTTGGAAATGGAAGATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 122

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38

Db 123 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGGATTCTGAGAGTCTCTGT 176

RESULT 6

ABV03916

ID ABV03916 standard; cDNA; 508 BP.

XX AC ABV03916;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 3907.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX PS Claim 1; Page 692; 11750pp; English.

XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (i) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 508 BP; 176 A; 100 C; 120 G; 111 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 4.39e-21 Length: 508
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x ABV03916 (1-508)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20

Db 58 AATAAGCCTTGGAAATGGAAGATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 117

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38

Db 118 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGGATTCTGAGAGTCTCTGT 171

RESULT 7

AAL16216

ID AAL16216 standard; cDNA; 533 BP.

XX AC AAL16216;

XX DT 07-DEC-2001 (first entry)

XX DE Human breast cancer expressed polynucleotide 8673.

XX KW Human; breast cancer; cell marker; cytostatic; ss.

XX OS Homo sapiens.

XX PN WO200151628-A2.

XX PD 19-JUL-2001.

XX PF 10-JAN-2001; 2001WO-US000798.

XX PR 14-JAN-2000; 2000US-0176077P.

XX PR 14-MAR-2000; 2000US-0189167P.

XX PR 24-MAR-2000; 2000US-0192099P.

XX PR 29-MAR-2000; 2000US-0193480P.

XX PR 15-MAY-2000; 2000US-0205230P.

XX PR 09-JUN-2000; 2000US-0211315P.

XX PR 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX DR WPI; 2001-451856/48.

XX PT New peptide useful as a marker for the diagnosis of breast cancer.

XX PS Claim 1; Page 1570; 3695pp; English.

XX CC The invention relates to human breast cancer expressed polynucleotides
XX (AAL07544-AAL26789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterising treating and
XX potentially preventing breast cancer. The polynucleotides and encoded

CC polypeptides are also useful for isolating compounds with cytostatic activity
XX
SQ Sequence 533 BP; 188 A; 106 C; 117 G; 122 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,67e-21 Length: 533
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAL16216 (1-533)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
DB 85 AATTAAGCCTTGGAATTGAGATGATGAACAACATTGAGAGCAGATGATCCCATCA 144
QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 145 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 198

RESULT 8

AAL25059
ID AAL25059 standard; cDNA; 562 BP.

XX AC AAL25059;

XX DT 07-DEC-2001 (first entry)

XX DE Human breast cancer expressed polynucleotide 17516.

XX KW Human; breast cancer; cell marker; cytostatic; ss.

XX OS Homo sapiens.

XX PN WO200151628-A2.

XX PD 19-JUL-2001.

XX PF 10-JAN-2001; 2001WO-US000798.

XX PR 14-JAN-2000; 2000US-0176077P.

XX PR 14-MAR-2000; 2000US-0189167P.

XX PR 24-MAR-2000; 2000US-0192099P.

XX PR 29-MAR-2000; 2000US-0193480P.

XX PR 15-MAY-2000; 2000US-0205230P.

XX PR 09-JUN-2000; 2000US-0211315P.

XX PR 25-JUL-2000; 2000US-0220534P.

XX (WILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX DR New peptide useful as a marker for the diagnosis of breast cancer.

XX PS Claim 1; Page 3237; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

XX SQ Sequence 562 BP; 196 A; 105 C; 126 G; 135 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5e-21 Length: 562
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAL25059 (1-562)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
DB 63 AATTAAGCCTTGGAATTGAGATGATGAACAACATTGAGAGCAGATGATCCCATCA 122
QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 123 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 176

RESULT 9

AAL20310/c

ID AAL20310 standard; cDNA; 572 BP.

XX AC AAL20310;

XX DT 07-DEC-2001 (first entry)

XX DE Human breast cancer expressed polynucleotide 12767.

XX KW Human; breast cancer; cell marker; cytostatic; ss.

XX OS Homo sapiens.

XX PN WO200151628-A2.

XX PD 19-JUL-2001.

XX PF 10-JAN-2001; 2001WO-US000798.

XX PR 14-JAN-2000; 2000US-0176077P.

XX PR 14-MAR-2000; 2000US-0189167P.

XX PR 24-MAR-2000; 2000US-0192099P.

XX PR 29-MAR-2000; 2000US-0193480P.

XX PR 15-MAY-2000; 2000US-0205230P.

XX PR 09-JUN-2000; 2000US-0211315P.

XX PR 25-JUL-2000; 2000US-0220534P.

XX (WILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX DR New peptide useful as a marker for the diagnosis of breast cancer.

XX PS Claim 1; Page 2256; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity

XX SQ Sequence 572 BP; 137 A; 128 C; 109 G; 198 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.12e-21 Length: 572
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0

```
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAL20310 (1-572)

Qy 1 AsnLySaLaLeuGluLeuLyAsnGluGlnThrLeuArgAlaAspGluIleLeuProSer 20
Db 510 AATAAAGCCTTGGATTGAGAAATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 451

Qy 21 GluSerLySgLnLyAspTyrgLluSerSerTrpAspSerGluSerIeuCys 38
Db 450 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGTGAGAGTCTGT 397

RESULT 10
AAL20029
ID AAL20029 standard; cDNA; 578 BP.
XX
AC AAL20029;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 12486.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
XX
PR 14-MAR-2000; 2000US-0189167P.
XX
PR 24-MAR-2000; 2000US-0192099P.
XX
PR 29-MAR-2000; 2000US-0193480P.
XX
PR 15-MAY-2000; 2000US-0205230P.
XX
PR 09-JUN-2000; 2000US-0211315P.
XX
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
WPI; 2001-451856/48.
XX
New peptide useful as a marker for the diagnosis of breast cancer.
XX
Claim 1; Page 2207-2208; 3695pp; English.
XX
The invention relates to human breast cancer expressed polynucleotides (AAL07544-AAL26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity
XX
SQ Sequence 578 BP; 201 A; 112 C; 124 G; 141 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,19e-21 Length: 578
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAL20029 (1-578)

Qy 1 AsnLySaLaLeuGluLeuLyAsnGluGlnThrLeuArgAlaAspGluIleLeuProSer 20
Db 510 AATAAAGCCTTGGATTGAGAAATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 451

Qy 21 GluSerLySgLnLyAspTyrgLluSerSerTrpAspSerGluSerIeuCys 38
Db 450 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGTGAGAGTCTGT 397

RESULT 11
AAH55638/c
ID AAH55638 standard; DNA; 580 BP.
XX
AC AAH55638;
XX
DT 04-SEP-2001 (first entry)
XX
DE Human breast tumour protein clone 48970 DNA sequence.
XX
KW Cytostatic; vaccine; human; breast tumour protein; breast cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200140269-A2.
XX
PD 07-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US032520.
XX
PR 30-NOV-1999; 99US-00451651.
XX
PR 22-FEB-2000; 2000US-00510662.
XX
PR 10-MAR-2000; 2000US-00523586.
XX
PR 07-APR-2000; 2000US-00545068.
XX
PR 15-MAY-2000; 2000US-00571025.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang A;
XX
WPI; 2001-356154/37.
XX
Breast tumor polypeptides and the nucleic acids that encode them, useful for the prevention, diagnosis and treatment of breast cancer.
XX
Claim 24; Page 172; 221pp; English.
XX
The present sequence is a human breast tumour protein coding sequence. This sequence may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of the breast tumour protein e.g. breast cancer. For example, this sequence may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of breast tumour protein by expressing inactive proteins or to supplement the patients own production of the breast tumour protein. Additionally, the present sequence may be used to produce the breast tumour protein, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. The present sequence and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy
XX
SQ Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 5,21e-21 Length: 580
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAH55638 (1-580)
```

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
ADN40408/c
ID ADN40408 standard; cDNA; 580 BP.
XX
AC ADN40408;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human breast cancer associated cDNA sequence #160.
XX
KW Human; breast cancer; T cell; tumour protein; antigen presenting cell;
KW immune response; CD4+; CD8+; cytostatic; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN US2004101899-A1.
XX
PD 27-MAY-2004.
XX
PF 13-NOV-2003; 2003US-00714389.
XX
PR 30-NOV-1999; 99US-00451651.
PR 22-FEB-2000; 2000US-00510662.
PR 10-MAR-2000; 2000US-00523586.
PR 07-APR-2000; 2000US-00545068.
PR 15-MAY-2000; 2000US-00571025.
PR 06-FEB-2001; 2001US-0078320.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;
PI McNeill PD;
XX
DR WPI; 2004-399689/37.
XX
PT New polynucleotides, useful for treating and diagnosing cancer,
PT particularly breast cancer by stimulating immune response in a patient
PT and inhibiting the development of cancer.
XX
PS Claim 1; SEQ ID NO 163; 151pp; English.
XX
CC The present invention relates to polynucleotide and polypeptide sequences
CC associated with breast cancer. Also disclosed are expression vectors
CC comprising the polynucleotide sequences of the invention operably linked
CC to an expression control sequence, host cells comprising the vector,
CC antibodies (or antigen binding fragments of antibodies) specifically
CC binding the polypeptides of the invention, fusion proteins comprising at
CC least one of the polypeptides, stimulating and/or expanding T cells
CC specific for a tumour protein. The polynucleotide sequences, polypeptide
CC sequences, and antigen presenting cells can be administered
CC therapeutically/prophylactically to induce an immune response. They can
CC be included with a physiological carrier/immunostimulant in compositions
CC such as vaccines, particularly to treat or prevent cancers such as breast
CC cancer. They can also be used to inhibit the development of cancer by
CC incubating one or more of them with CD4+ and/or CD8+ T cells isolated
CC from a patient, such that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The polynucleotide sequences are
CC useful for detecting cancer in a patient, producing fusion proteins,
CC producing T cell populations and antigen presenting cells. The present
CC sequence represents a polynucleotide sequence of the invention.
XX
SQ Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;
XX
Alignment Scores: 5.21e-21 Length: 580
Pred. No.:

Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-09-489-079-27 (1-38) x ADN40408 (1-580)
QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
ADN40408/c
ID ADN40408 standard; cDNA; 580 BP.
XX
AC ADN40408;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human breast cancer associated cDNA sequence #160.
XX
KW Human; breast cancer; T cell; tumour protein; antigen presenting cell;
KW immune response; CD4+; CD8+; cytostatic; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN US2004101899-A1.
XX
PD 27-MAY-2004.
XX
PF 13-NOV-2003; 2003US-00714389.
XX
PR 30-NOV-1999; 99US-00451651.
PR 22-FEB-2000; 2000US-00510662.
PR 10-MAR-2000; 2000US-00523586.
PR 07-APR-2000; 2000US-00545068.
PR 15-MAY-2000; 2000US-00571025.
PR 06-FEB-2001; 2001US-0078320.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;
PI McNeill PD;
XX
DR WPI; 2004-399689/37.
XX
PT New polynucleotides, useful for treating and diagnosing cancer,
PT particularly breast cancer by stimulating immune response in a patient
PT and inhibiting the development of cancer.
XX
PS Claim 1; SEQ ID NO 163; 151pp; English.
XX
CC The present invention relates to polynucleotide and polypeptide sequences
CC associated with breast cancer. Also disclosed are expression vectors
CC comprising the polynucleotide sequences of the invention operably linked
CC to an expression control sequence, host cells comprising the vector,
CC antibodies (or antigen binding fragments of antibodies) specifically
CC binding the polypeptides of the invention, fusion proteins comprising at
CC least one of the polypeptides, stimulating and/or expanding T cells
CC specific for a tumour protein. The polynucleotide sequences, polypeptide
CC sequences, and antigen presenting cells can be administered
CC therapeutically/prophylactically to induce an immune response. They can
CC be included with a physiological carrier/immunostimulant in compositions
CC such as vaccines, particularly to treat or prevent cancers such as breast
CC cancer. They can also be used to inhibit the development of cancer by
CC incubating one or more of them with CD4+ and/or CD8+ T cells isolated
CC from a patient, such that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The polynucleotide sequences are
CC useful for detecting cancer in a patient, producing fusion proteins,
CC producing T cell populations and antigen presenting cells. The present
CC sequence represents a polynucleotide sequence of the invention.
XX
SQ Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;
XX
Alignment Scores: 5.21e-21 Length: 580
Pred. No.:

Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-09-489-079-27 (1-38) x ADN40408 (1-580)
QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
ADN40408/c
ID ADN40408 standard; cDNA; 580 BP.
XX
AC ADN40408;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human breast cancer associated cDNA sequence #160.
XX
KW Human; breast cancer; T cell; tumour protein; antigen presenting cell;
KW immune response; CD4+; CD8+; cytostatic; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN US2004101899-A1.
XX
PD 27-MAY-2004.
XX
PF 13-NOV-2003; 2003US-00714389.
XX
PR 30-NOV-1999; 99US-00451651.
PR 22-FEB-2000; 2000US-00510662.
PR 10-MAR-2000; 2000US-00523586.
PR 07-APR-2000; 2000US-00545068.
PR 15-MAY-2000; 2000US-00571025.
PR 06-FEB-2001; 2001US-0078320.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;
PI McNeill PD;
XX
DR WPI; 2004-399689/37.
XX
PT New polynucleotides, useful for treating and diagnosing cancer,
PT particularly breast cancer by stimulating immune response in a patient
PT and inhibiting the development of cancer.
XX
PS Claim 1; SEQ ID NO 163; 151pp; English.
XX
CC The present invention relates to polynucleotide and polypeptide sequences
CC associated with breast cancer. Also disclosed are expression vectors
CC comprising the polynucleotide sequences of the invention operably linked
CC to an expression control sequence, host cells comprising the vector,
CC antibodies (or antigen binding fragments of antibodies) specifically
CC binding the polypeptides of the invention, fusion proteins comprising at
CC least one of the polypeptides, stimulating and/or expanding T cells
CC specific for a tumour protein. The polynucleotide sequences, polypeptide
CC sequences, and antigen presenting cells can be administered
CC therapeutically/prophylactically to induce an immune response. They can
CC be included with a physiological carrier/immunostimulant in compositions
CC such as vaccines, particularly to treat or prevent cancers such as breast
CC cancer. They can also be used to inhibit the development of cancer by
CC incubating one or more of them with CD4+ and/or CD8+ T cells isolated
CC from a patient, such that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The polynucleotide sequences are
CC useful for detecting cancer in a patient, producing fusion proteins,
CC producing T cell populations and antigen presenting cells. The present
CC sequence represents a polynucleotide sequence of the invention.
XX
SQ Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;
XX
Alignment Scores: 5.21e-21 Length: 580
Pred. No.:

Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-09-489-079-27 (1-38) x ADN40408 (1-580)
QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
ADN40408/c
ID ADN40408 standard; cDNA; 580 BP.
XX
AC ADN40408;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human breast cancer associated cDNA sequence #160.
XX
KW Human; breast cancer; T cell; tumour protein; antigen presenting cell;
KW immune response; CD4+; CD8+; cytostatic; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN US2004101899-A1.
XX
PD 27-MAY-2004.
XX
PF 13-NOV-2003; 2003US-00714389.
XX
PR 30-NOV-1999; 99US-00451651.
PR 22-FEB-2000; 2000US-00510662.
PR 10-MAR-2000; 2000US-00523586.
PR 07-APR-2000; 2000US-00545068.
PR 15-MAY-2000; 2000US-00571025.
PR 06-FEB-2001; 2001US-0078320.
XX
PA (CORI-) CORIXA CORP.
XX
PI Dillon DC, Day CH, Jiang Y, Houghton RL, Mitcham JL, Wang T;
PI McNeill PD;
XX
DR WPI; 2004-399689/37.
XX
PT New polynucleotides, useful for treating and diagnosing cancer,
PT particularly breast cancer by stimulating immune response in a patient
PT and inhibiting the development of cancer.
XX
PS Claim 1; SEQ ID NO 163; 151pp; English.
XX
CC The present invention relates to polynucleotide and polypeptide sequences
CC associated with breast cancer. Also disclosed are expression vectors
CC comprising the polynucleotide sequences of the invention operably linked
CC to an expression control sequence, host cells comprising the vector,
CC antibodies (or antigen binding fragments of antibodies) specifically
CC binding the polypeptides of the invention, fusion proteins comprising at
CC least one of the polypeptides, stimulating and/or expanding T cells
CC specific for a tumour protein. The polynucleotide sequences, polypeptide
CC sequences, and antigen presenting cells can be administered
CC therapeutically/prophylactically to induce an immune response. They can
CC be included with a physiological carrier/immunostimulant in compositions
CC such as vaccines, particularly to treat or prevent cancers such as breast
CC cancer. They can also be used to inhibit the development of cancer by
CC incubating one or more of them with CD4+ and/or CD8+ T cells isolated
CC from a patient, such that the T cells proliferate, and administering the
CC proliferated T cells to the patient. The polynucleotide sequences are
CC useful for detecting cancer in a patient, producing fusion proteins,
CC producing T cell populations and antigen presenting cells. The present
CC sequence represents a polynucleotide sequence of the invention.
XX
SQ Sequence 580 BP; 139 A; 127 C; 112 G; 199 T; 0 U; 3 Other;
XX
Alignment Scores: 5.21e-21 Length: 580
Pred. No.:

US-09-489-079-27 (1-38) x ACN92774 (1-600)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
 Db 454 AATAAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCAGATGAGATATCTCCATCA 513

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTTPAspSerGluSerLeuCys 38
 Db 514 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 567

RESULT 14

ACN87508
 ID ACN87508 standard; DNA; 635 BP.

XX ACN87508;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 8658.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for
 PT detecting presence of polypeptide in sample, as a marker for breast
 PT cancer.

XX Disclosure; SEQ ID NO 8658; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with a
 CC breast cancer which is encoded by a nucleic acid molecule comprising a
 CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
 CC the polypeptide of the invention. The activity of the polypeptide of the
 CC invention may be described as cytostatic. The antibody is useful for
 CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
 CC invention are useful in the detection of breast tumours. (I) is useful as
 CC a marker for breast cancer and in breast cancer therapy. Sequences given
 CC in records ACN78851-ACN92934 represent nucleic acid markers associated
 CC with breast cancer. Note: The sequence listing does not form part of the
 CC specification but may be obtained in electronic format from the USPTO web
 CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX Sequence 635 BP; 213 A; 128 C; 154 G; 135 T; 0 U; 5 Other;

Alignment Scores:
 Pred. No.: 5,86e-21 Length: 635
 Score: 195.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 11 Gaps: 0

US-09-489-079-27 (1-38) x ACN87508 (1-635)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
 Db 396 AATAAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCAGATGAGATATCTCCATCA 455

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTTPAspSerGluSerLeuCys 38
 Db 456 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 509

RESULT 15

AAF22981
 ID AAF22981 standard; cDNA; 636 BP.

XX AAF22981;

XX 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen nucleotide sequence SEQ ID:560.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine; ss.

XX Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014749.

XX 28-MAY-1999; 99US-0136526P.

PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.

XX Claim 50; Page 456; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer

XX Sequence 636 BP; 217 A; 120 C; 138 G; 157 T; 0 U; 4 Other;

Alignment Scores:
 Pred. No.: 5,87e-21 Length: 636
 Score: 195.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0

US-09-489-079-27 (1-38) x AAF22981 (1-636)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
 Db 252 AATAAAGCCTTGGAAATGAAGAAATGAACAAACATTGAGAGCAGATGAGATATCTCCATCA 311

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTTPAspSerGluSerLeuCys 38

Db 312 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 365

Search completed: August 1, 2005, 23:18:12
Job time : 71.372 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:24:58 ; Search time 575.313 Seconds
(without alignments)
3200.517 Million cell updates/sec

Title: US-09-489-079-27

Perfect score: 195

Sequence: 1 NKAELKNEQTLRADEILPSESKQDYESSWDSRLC 38

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 segs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09489079/runat_01082005_141250_13297/app_query.fasta_1.1635
-DB=GenEmbl -QFMT=fastap -SUFFIX=erge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079_@CNG_1_1_8858 @runat_01082005_141250_13297 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBIOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.cm.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.ty.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	278	6	BD271305
2	195	100.0	445	6	CQ419026
3	195	100.0	466	6	CQ481209
4	195	100.0	473	6	CQ502357

5	195	100.0	473	6	CQ511219
6	195	100.0	508	6	CQ472041
7	195	100.0	533	6	CQ423647
8	195	100.0	562	6	CQ432495
9	195	100.0	572	6	CQ427744
10	195	100.0	578	6	CQ427463
11	195	100.0	580	6	AX156306
12	195	100.0	636	6	AX053794
13	195	100.0	664	6	AX053793
14	195	100.0	708	6	AX053785
15	195	100.0	713	6	CQ417876
16	195	100.0	713	6	AX053787
17	195	100.0	741	6	AX053804
18	195	100.0	766	6	AX053799
19	195	100.0	770	6	CQ427927
20	195	100.0	771	6	AX053802
21	195	100.0	776	6	AX053782
22	195	100.0	814	6	AX053780
23	195	100.0	833	6	AX053800
24	195	100.0	904	6	AX053778
25	195	100.0	1158	6	AX053798
26	195	100.0	1337	6	AR283455
27	195	100.0	1337	6	AR344223
28	195	100.0	1337	6	AR351424
29	195	100.0	1337	6	AR454004
30	195	100.0	1337	6	AR561592
31	195	100.0	1337	6	AX282974
32	195	100.0	1337	6	AX303147
33	195	100.0	1448	6	CQ490684
34	195	100.0	1448	6	CQ495497
35	195	100.0	1448	6	CQ496522
36	195	100.0	1729	6	AR283454
37	195	100.0	1729	6	AR344222
38	195	100.0	1729	6	AR351423
39	195	100.0	1729	6	AR454003
40	195	100.0	1729	6	AR561591
41	195	100.0	1729	6	AX282973
42	195	100.0	1729	6	AX303146
43	195	100.0	2030	6	AX367056
44	195	100.0	2232	6	AR283468
45	195	100.0	2232	6	AR454017

ALIGNMENTS

RESULT 1
BD271305
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD271305 278 bp DNA linear PAT 17-JUL-2003
Reagents and methods useful for detecting diseases of the breast.

BD271305.1 GI:33081073
JP 2002540761-A/2.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

Medel, P.A.B., Cohen, M., Colpitts, T.L., Friedman, P.N., Gordon, J.,

Granados, E.N., Hodges, S.C., Klass, M.R., Kratochvil, J.D.,

Russell, J.C. and Stroupe, S.D.

Reagents and methods useful for detecting diseases of the breast

Patent: JP 2002540761-A 2 03-DEC-2002;

ABBOTT LABORATORIES

OS Homo sapiens (human)

PN JP 2002540761-A/2

PD 03-DEC-2002

PF 21-JAN-2000 JP 2000594836

PR 21-JAN-1999 US 09/234716

PI PATRICIA A BILLING MEDEL, MAURICE COHEN, TRACEY L COLPITTS, PAULA

PI N FRIEDMAN,

PI JULIAN GORDON, EDWARD N GRANADOS, STEVEN C HODGES, MICHAEL R PI

KLASS,

PI JON D KRATOCHVIL, JOHN C RUSSELL, STEPHEN D STROUPE PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10, C12P21/02, C12Q1/68, G01N33/53, G01N33/574, G01N33/574,
PC G01N37/00,
PC C12N15/00, C12N5/00
CC Reagents and methods useful for detecting diseases of the CC
FH Key Location/Qualifiers
FT source 1..278 /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..278 /organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 2,15e-22 Length: 278
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x BD271305 (1-278)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
|||||
Db 22 AATAAAGCCTTGAATTTGAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 81
|||||

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
|||||
Db 82 GAATCCAAACAAAGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 135
|||||

RESULT 2
CQ419026
LOCUS
DEFINITION
Sequence 4060 from Patent WO0151628.
ACCESSION
CQ419026
VERSION
CQ419026.1 GI:41371255
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE
Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL
Patent: WO 0151628-A 4060 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..445 /organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 3,61e-22 Length: 445
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x CQ419026 (1-445)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
|||||
Db 57 AATAAAGCCTTGAATTTGAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 116
|||||

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
|||||
Db 117 GAATCCAAACAAAGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 170
|||||

RESULT 3
CQ481209
LOCUS
DEFINITION
Sequence 13076 from Patent WO0160860.
ACCESSION
CQ481209
VERSION
CQ481209.1 GI:41446828
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE
Genes differentially expressed in human prostate cancer and their
use
JOURNAL
Patent: WO 0160860-A 13076 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source Location/Qualifiers
1..466 /organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 3,8e-22 Length: 466
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x CQ481209 (1-466)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
|||||
Db 21 AATAAAGCCTTGAATTTGAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 80
|||||

Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
|||||
Db 81 GAATCCAAACAAAGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 134
|||||

RESULT 4
CQ502357
LOCUS
DEFINITION
Sequence 34224 from Patent WO0160860.
ACCESSION
CQ502357
VERSION
CQ502357.1 GI:41467993
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE
Genes differentially expressed in human prostate cancer and their
use
JOURNAL
Patent: WO 0160860-A 34224 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source Location/Qualifiers
1..473 /organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'

ORIGIN
Alignment Scores:
Pred. No.: 3,86e-22 Length: 473
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x CQ502357 (1-473)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluIleLeuProSer 20
Db 63 AATAAAGCCTTGGAAATTAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 122

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 123 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAGAGTCTCTGT 176

RESULT 5
CQ511219 473 bp DNA linear PAT 30-JAN-2004
LOCUS
DEFINITION Sequence 43086 from Patent WO0160860.
ACCESSION CQ511219
VERSION CQ511219.1 GI:41477483
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 43086 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES
source
1..473
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 3.86e-22 Length: 473
Pred. No.: 195.00 Matches: 38
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-489-079-27 (1-38) x CQ511219 (1-473)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluIleLeuProSer 20
Db 63 AATAAAGCCTTGGAAATTAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 122

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 123 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAGAGTCTCTGT 176

RESULT 6
CQ472041 508 bp DNA linear PAT 30-JAN-2004
LOCUS
DEFINITION Sequence 3907 from Patent WO0160860.
ACCESSION CQ472041
VERSION CQ472041.1 GI:41437660
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use

JOURNAL Patent: WO 0160860-A 3907 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES
source
1..508
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 4.18e-22 Length: 508
Pred. No.: 195.00 Matches: 38
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-489-079-27 (1-38) x CQ472041 (1-508)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluIleLeuProSer 20
Db 58 AATAAAGCCTTGGAAATTAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 117

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 118 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAGAGTCTCTGT 171

RESULT 7
CQ423647 533 bp DNA linear PAT 28-JAN-2004
LOCUS
DEFINITION Sequence 8681 from Patent WO0151628.
ACCESSION CQ423647
VERSION CQ423647.1 GI:41375876
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
AUTHORS Lillie, J., Xu, Y., Wang, Y. and Steinmann, K.
TITLE Novel genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 8681 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES
source
1..533
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 4.4e-22 Length: 533
Pred. No.: 195.00 Matches: 38
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-489-079-27 (1-38) x CQ423647 (1-533)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluIleLeuProSer 20
Db 85 AATAAAGCCTTGGAAATTAAGAAATGAACAACATTGAGAGCAGATGAGATCTCCATCA 144

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 145 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTTCTGAGAGTCTCTGT 198

RESULT 8
CQ432495 562 bp DNA linear PAT 28-JAN-2004
LOCUS
DEFINITION Sequence 17529 from Patent WO0151628.

```

ACCESSION   CQ432495
VERSION     CQ432495.1  GI:41384724
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE       Novel genes, compositions, kits, and methods for identification,
JOURNAL     assessment, prevention, and therapy of breast cancer
            Patent: WO 0151628-A 17529 19-JUL-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES   source
            Location/Qualifiers
            1..562
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      4,678-22      Length:      562
Score:          195.00      Matches:    38
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0

US-09-489-079-27 (1-38) x CQ432495 (1-562)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
Db 63 AATAAAGCCTTGAATTGAAGATGAACAAACATTGAGAGCAGATGAGTACTCCATCA 122
Qy 21 GluSerLysGlnLysAspTyrGluGluSerTrpAspSerGluSerLeuCys 38
Db 123 GAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCGTGAGAGTCTCTGT 176

RESULT 9
CQ427744/c
LOCUS       CQ427744      572 bp      DNA      linear      PAT 28-JAN-2004
DEFINITION  Sequence 12778 from Patent WO0151628.
ACCESSION   CQ427744
VERSION     CQ427744.1  GI:41379973
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE       Novel genes, compositions, kits, and methods for identification,
JOURNAL     assessment, prevention, and therapy of breast cancer
            Patent: WO 0151628-A 12778 19-JUL-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES   source
            Location/Qualifiers
            1..572
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      4,768-22      Length:      572
Score:          195.00      Matches:    38
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0

US-09-489-079-27 (1-38) x CQ427744 (1-572)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
Db 75 AATAAAGCCTTGAATTGAAGATGAACAAACATTGAGAGCAGATGAGTACTCCATCA 134
Qy 21 GluSerLysGlnLysAspTyrGluGluSerTrpAspSerGluSerLeuCys 38
Db 135 GAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCGTGAGAGCCTCTGT 188

RESULT 11
AX156306/c
LOCUS       AX156306      580 bp      DNA      linear      PAT 22-JUN-2001
DEFINITION  Sequence 163 from Patent WO0140269.
ACCESSION   AX156306
VERSION     AX156306.1  GI:14537307
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Dillon,D.C., Day,C.H., Jiang,Y., Houghton,R.L., Mitcham,J.L. and
            Wang,A.
TITLE       Compositions and methods for therapy and diagnosis of breast cancer
JOURNAL     Patent: WO 0140269-A 163 07-JUN-2001;
            CORIXA CORPORATION (US)
FEATURES   source
            Location/Qualifiers
            1..580
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      4,828-22      Length:      578
Score:          195.00      Matches:    38
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0

US-09-489-079-27 (1-38) x CQ427463 (1-578)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
Db 75 AATAAAGCCTTGAATTGAAGATGAACAAACATTGAGAGCAGATGAGTACTCCATCA 134
Qy 21 GluSerLysGlnLysAspTyrGluGluSerTrpAspSerGluSerLeuCys 38
Db 135 GAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCGTGAGAGCCTCTGT 188

RESULT 10
CQ427463
LOCUS       CQ427463      578 bp      DNA      linear      PAT 28-JAN-2004
DEFINITION  Sequence 12497 from Patent WO0151628.
ACCESSION   CQ427463
VERSION     CQ427463.1  GI:41379692
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE       Novel genes, compositions, kits, and methods for identification,
JOURNAL     assessment, prevention, and therapy of breast cancer
            Patent: WO 0151628-A 12497 19-JUL-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES   source
            Location/Qualifiers
            1..578
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      4,828-22      Length:      578
Score:          195.00      Matches:    38
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:    0
DB:             6           Gaps:      0

US-09-489-079-27 (1-38) x CQ427463 (1-578)

Qy 1 AsnLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
Db 75 AATAAAGCCTTGAATTGAAGATGAACAAACATTGAGAGCAGATGAGTACTCCATCA 134
Qy 21 GluSerLysGlnLysAspTyrGluGluSerTrpAspSerGluSerLeuCys 38
Db 135 GAATCCAAACAAAGGACTATGAGAAAGTTCTTGGGATTCGTGAGAGCCTCTGT 188

RESULT 11
AX156306/c
LOCUS       AX156306      580 bp      DNA      linear      PAT 22-JUN-2001
DEFINITION  Sequence 163 from Patent WO0140269.
ACCESSION   AX156306
VERSION     AX156306.1  GI:14537307
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Dillon,D.C., Day,C.H., Jiang,Y., Houghton,R.L., Mitcham,J.L. and
            Wang,A.
TITLE       Compositions and methods for therapy and diagnosis of breast cancer
JOURNAL     Patent: WO 0140269-A 163 07-JUN-2001;
            CORIXA CORPORATION (US)
FEATURES   source
            Location/Qualifiers
            1..580
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN

```

```

Alignment Scores:
Pred. No.: 4,84e-22 Length: 580
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x AX156306 (1-580)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
DB 505 AATAAGCCTTGGAAATGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 446

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 445 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 392

RESULT 12
AX053794
LOCUS AX053794 636 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 560 from Patent WO0073801.
ACCESSION AX053794
VERSION AX053794.1 GI:12228113
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Obata,Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
JOURNAL therefor
PATENT: WO 0073801-A 560 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5,35e-22 Length: 636
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x AX053794 (1-636)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
DB 252 AATAAGCCTTGGAAATGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 311

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 312 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 365

RESULT 13
AX053793
LOCUS AX053793 664 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 559 from Patent WO0073801.
ACCESSION AX053793
VERSION AX053793.1 GI:12228112
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Obata,Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
JOURNAL therefor
PATENT: WO 0073801-A 551 07-DEC-2000;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 6,03e-22 Length: 708
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x AX053785 (1-708)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
DB 289 AATAAGCCTTGGAAATGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 348

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 349 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 402

RESULT 15
CQ417876

```

```

TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor
JOURNAL LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 5,61e-22 Length: 664
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x AX053793 (1-664)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
DB 42 AATAAGCCTTGGAAATGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 101

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 102 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 155

RESULT 14
AX053785
LOCUS AX053785 708 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 551 from Patent WO0073801.
ACCESSION AX053785
VERSION AX053785.1 GI:12228104
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Obata,Y.
TITLE Breast, gastric and prostate cancer associated antigens and uses
therefor
JOURNAL LUDWIG INSTITUTE FOR CANCER RESEARCH (US)
FEATURES
source
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 6,03e-22 Length: 708
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-489-079-27 (1-38) x AX053785 (1-708)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
DB 289 AATAAGCCTTGGAAATGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 348

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
DB 349 GAATCCAAACAAGGACTATGAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 402

RESULT 15
CQ417876

```

LOCUS CQ417876 713 bp DNA linear PAT 28-JAN-2004
DEFINITION Sequence 2910 from Patent WO0151628.
ACCESSION CQ417876
VERSION CQ417876.1 GI:41370105
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE Novel genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of breast cancer
JOURNAL Patent: WO 0151628-A 2910 19-JUL-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..713
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.07e-22 Length: 713
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-489-079-27 (1-38) x CQ417876 (1-713)
Qy 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeuProSer 20
Db 118 AATAAAGCCTTGGAATTGAAGAATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 177
Qy 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
Db 178 GAATCAACAACAAAGGACTATGAGAAAGTTCTTGGGATTCGAGAGCCTCTGT 231

Search completed: August 2, 2005, 02:51:22
Job time : 577.313 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:08:06 ; Search time 2.34123 Seconds
(without alignments)
1561.672 Million cell updates/sec

Title: US-09-489-079-27

Perfect score: 195

Sequence: 1 NKALELKNQTLRADELIPSESKQDYESSWDSBLC 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	58.5	30.0	678	2 T32483	hypothetical prote
2	56.5	29.0	707	2 F83465	flagellar biosynth
3	56	28.7	752	2 S51866	HPRI protein - yea
4	55.5	28.5	237	2 A85901	probable alpha hel
5	55.5	28.5	237	2 A49940	probable alpha hel
6	55.5	28.5	237	2 E91056	probable alpha hel
7	55.5	28.5	332	2 C40646	endospore developm
8	54	27.7	173	2 T47279	hypothetical prote
9	54	27.7	396	2 D64509	hypothetical prote
10	54	27.7	1403	1 A47328	natural killer cel
11	53.5	27.4	511	2 T45869	hypothetical prote
12	53	27.2	375	2 T32251	hypothetical prote
13	52.5	26.9	494	2 E83254	glutamyl-tRNA synt
14	52.5	26.9	629	2 T15945	hypothetical prote
15	52.5	26.9	784	1 A55236	kinesin-related pr
16	52.5	26.9	2331	2 T25410	hypothetical prote
17	52	26.7	139	2 T46457	hypothetical prote
18	52	26.7	237	2 S34797	listeriolysin regu
19	52	26.7	257	2 T12092	G-box-binding prot
20	52	26.7	616	2 E85508	hypothetical prote
21	52	26.7	616	2 E90657	hypothetical prote
22	52	26.7	697	2 JN0602	rfb protein - Esch
23	52	26.7	876	2 G90592	hypothetical prote
24	52	26.7	1080	2 T00587	probable ubiquitin
25	51.5	26.4	232	2 G96585	hypothetical prote
26	51.5	26.4	432	2 T34154	hypothetical prote
27	51.5	26.4	569	1 VHXPLJ	major structural n
28	51.5	26.4	720	2 G64230	stringent response
29	51.5	26.4	1674	2 T01265	starch synthase DU

30 51 26.2 257 2 G90727 molybdate-binding
31 51 26.2 257 2 H85578 molybdate-binding
32 51 26.2 356 2 T06748 hypothetical prote
33 51 26.2 492 2 H72473 probable phosphosu
34 51 26.2 521 2 F82918 ammonium transport
35 51 26.2 597 2 T45676 hypothetical prote
36 51 26.2 793 2 JC5831 kinesin-related pr
37 51 26.2 1104 2 C72409 reverse gyrase - T
38 51 26.2 1856 2 B35049 ankryrin 1, erythro
39 51 26.2 1880 2 A35049 ankryrin 1, erythro
40 51 26.2 1881 1 SJHUK ankryrin 1, erythro
41 50.5 25.9 371 1 F69991 N-acylamino acid r
42 50.5 25.9 511 2 S24345 Balbiani ring 1 pr
43 50.5 25.9 522 2 G86259 myb-related protei
44 50.5 25.9 590 1 S55095
45 50.5 25.9 1228 2 S46754 hypothetical prote

ALIGNMENTS

RESULT 1

T32483

hypothetical protein F52H2.1 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C;Accession: T32483

R;Chisoe, S.; Hawkins, J.

submitted to the EMBL Data Library, September 1997

A;Description: The sequence of C. elegans cosmid F52H2.

A;Reference number: Z21175

A;Accession: T32483

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-678 <CHI>

A;Cross-references: UNIPROT:O17399; EMBL:AF026214; PIDN:AAB71311.1; GSPDB:GN00028; CESP

A;Experimental source: strain Bristol N2; clone F52H2

C;Genetics:

A;Gene: CRSP:F52H2.1

A;Map position: X

A;Introns: 53/3; 96/3; 132/3; 169/3; 249/2; 373/3; 638/1

C;Superfamily: Caenorhabditis elegans hypothetical protein F52H2.1

Query Match 30.0%; Score 58.5; DB 2; Length 678;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 15; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

QY 1 NKALELKNQTLRADELIPSESKQDYESS 30

DB 624 NKALELKNQTLRADELIPSESKQDYESS 652

RESULT 2

F83465

flagellar biosynthesis protein FlhA PA1452 [imported] - Pseudomonas aeruginosa (strain

C;Species: Pseudomonas aeruginosa

C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: F83465

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A;Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: F83465

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-707 <STO>

A;Cross-references: UNIPROT:Q9I3P9; GB:AE004574; GB:AE004091; NID:G9947391; PIDN:AAG048

A;Experimental source: strain PA01

C;Genetics:

A;Gene: flhA; PA1452

C;Superfamily: regulatory protein lcrd

A; Residues: 1-175 <NTR>
A; Cross-references: UNIPROT:Q92233; EMBL:AJ138645
A; Experimental source: cultivar Columbea; EAC clone F26B15
C; Genetics:
A; Map position: 3
A; Introns: 64/2
A; Note: F26B15.10
C; Superfamily: Arabidopsis thaliana hypothetical protein F26B15.10

Query Match	27.7%	Score 54;	DB 1;	Length 1403;
Best Local Similarity	58.8%	Pred. No. 1.6e+02;		
Matches	10;	Conservative	3;	Mismatches 4; Indels 0; Gaps 0;
QY	20	SESQKQDYESSWDS	36	
		: : : :		
Db	927	SKPKKNYAGSKWDS	943	

Query Match 26.9%; Score 52.5; DB 1; Length 784;
Best Local Similarity 41.9%; Pred. No. 1.3e+02;
Matches 13; Conservative 5; Mismatches 12; Indels 1; Gaps 1;

Search completed: August 1, 2005, 22:23:20
Job time : 4.34123 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:44:18 ; Search time 1445.31 Seconds
(without alignments)
1000.786 Million cell updates/sec

Title: US-09-489-079-27

Perfect score: 195

Sequence: 1 NKAELKNEQTLRADEILFESKQKDYBESSWDSRLC 38

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US09489079/runat_01082005_141250_13309/app_query.fasta_1.1635
-DB=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=otc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CNG_1_1 10973 @runat_01082005_141250_13309 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	581	5	BP314867
2	195	100.0	582	5	BP313026
3	195	100.0	582	5	BP313235
4	195	100.0	582	5	BP313436
5	195	100.0	582	5	BP313704
6	195	100.0	582	5	BP314260
7	195	100.0	582	5	BP315806
8	195	100.0	583	5	BP312890
9	195	100.0	583	5	BP313822

10	195	100.0	584	5	BP314652	
11	195	100.0	598	5	BP312910	
12	182	93.3	443	2	BF746270	
13	182	93.3	443	2	BF746340	
14	180	92.3	583	5	BP313800	
15	169	86.7	582	5	BP315089	
C	16	160	82.1	490	1	AI951118
17	153	78.5	3443	3	BC028407	
18	145	74.4	574	2	BF328582	
C	19	122	62.6	563	8	AQ546028
C	20	93	47.7	705	8	AQ423531
21	92.5	47.4	344	8	AZ576125	
22	86	44.1	165	9	HSMC18H11	
23	85	43.6	155	2	BF171278	
24	74	37.9	431	2	BE062233	
25	67.5	34.6	535	5	BU811432	
26	67.5	34.6	550	7	CK102592	
27	67.5	34.6	553	4	B1128959	
C	28	67.5	34.6	847	1	AU132223
C	29	67.5	34.6	5193	3	CR627244
30	67	34.4	514	5	BP312630	
31	67	34.4	655	7	CN787741	
32	66	33.8	578	5	BP349185	
33	65.5	33.6	963	2	BF179023	
C	34	65.5	33.6	1069	9	CL109041
C	35	65.5	33.6	1075	9	CL108568
36	65	33.3	349	9	CE315218	
C	37	65	33.3	530	8	AZ110334
38	65	33.3	3527	3	HSME01761	
C	39	64	32.8	432	9	CE313306
40	63	32.3	389	7	CR735355	
41	63	32.3	596	5	BQ429618	
C	42	62	31.8	433	8	AQ090800
43	62	31.8	633	1	AV835394	
44	62	31.8	732	7	CN841887	
45	62	31.8	735	7	CK603996	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AGENCOURT	
					AG	

ALIGNMENTS

RESULT 1
BP314867
LOCUS
DEFINITION
BP314867 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
CDNA clone OFR07276, mRNA sequence.
581 bp mRNA linear EST 17-SEP-2004

ACCESSION
BP314867
VERSION
BP314867.1 GI:52243842
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 581)
AUTHORS
Suzuki,Y., Yanashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE
Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL
Genome Res. 14 (9), 1711-1718 (2004)
COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR07276"
/tissue type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

```

ORIGIN
Alignment Scores:
Pred. No.: 7.58e-18 Length: 581
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP314867 (1-581)
QY 1 AenLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
    |||||
DB 61 AATAAGCCTTGGATTGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 120
    |||||

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
    |||||
DB 121 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 174
    |||||

RESULT 2
BP313026
LOCUS
DEFINITION BP313026 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION BP313026
VERSION BP313026.1 GI:52242001
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
CONTACT: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR01939"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.: 7.59e-18 Length: 582
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP313235 (1-582)
QY 1 AenLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
    |||||
DB 262 AATAAGCCTTGGATTGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 321
    |||||

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
    |||||
DB 322 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 375
    |||||

RESULT 3
BP313235
LOCUS
DEFINITION BP313235 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION BP313235
VERSION BP313235.1 GI:52242210
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
CONTACT: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR02569"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.: 7.59e-18 Length: 582
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP313235 (1-582)
QY 1 AenLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuProSer 20
    |||||
DB 262 AATAAGCCTTGGATTGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 321
    |||||

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
    |||||
DB 322 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGATTCTGAGAGTCTCTGT 375
    |||||

RESULT 4
BP313436
LOCUS
DEFINITION BP313436 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
ACCESSION BP313436
VERSION BP313436.1 GI:52242411
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
CONTACT: Yutaka Suzuki
Department of Virology

```


Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasuku@ims.u-tokyo.ac.jp.

FEATURES

source

1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR03209"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 7,59e-18 Length: 582
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP313436 (1-582)

QY 1 AsnLYeAlaLeuGluLeuLYsAsnGluInThrLeuArgAlaAspGluLeuLeuProSer 20
|||||
DB 32 AATAAGCCCTTGAATTGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 91
|||||
QY 21 GluSerLYsGlnLYsAspTYrGluGluSerSerTrpAspSerGluSerLeuCys 38
|||||
DB 92 GAATCCAAACAAAGGACTATGAAGAAGTTCTTGGATTCTGAGAGTCTCTGT 145
|||||

RESULT 5

BP313704
LOCUS BP313704 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION CDNA clone OFR03974, mRNA sequence.

ACCESSION BP313704.1 GI:52242679

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasuku@ims.u-tokyo.ac.jp.

FEATURES

source

1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR03974"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 7,59e-18 Length: 582
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match:

DB: 100.00% Indels: 0

US-09-489-079-27 (1-38) x BP313704 (1-582)

QY 1 AsnLYeAlaLeuGluLeuLYsAsnGluInThrLeuArgAlaAspGluLeuLeuProSer 20
|||||

DB 57 AATAAGCCCTTGAATTGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 116
|||||

QY 21 GluSerLYsGlnLYsAspTYrGluGluSerSerTrpAspSerGluSerLeuCys 38
|||||

DB 117 GAATCCAAACAAAGGACTATGAAGAAGTTCTTGGATTCTGAGAGTCTCTGT 170
|||||

RESULT 6

LOCUS

BP314260
DEFINITION BP314260 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
CDNA clone OFR05703, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yasuku@ims.u-tokyo.ac.jp.

FEATURES

source

1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR05703"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN

Alignment Scores:

Pred. No.: 7,59e-18 Length: 582
Score: 195.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP314260 (1-582)

QY 1 AsnLYeAlaLeuGluLeuLYsAsnGluInThrLeuArgAlaAspGluLeuLeuProSer 20
|||||

DB 264 AATAAGCCCTTGAATTGAAGATGAACAACATTGAGAGCAGATGAGATCTCCATCA 323
|||||

QY 21 GluSerLYsGlnLYsAspTYrGluGluSerSerTrpAspSerGluSerLeuCys 38
|||||

DB 324 GAATCCAAACAAAGGACTATGAAGAAGTTCTTGGATTCTGAGAGTCTCTGT 377
|||||

RESULT 7

LOCUS

BP315806
DEFINITION BP315806 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
CDNA clone OFR09418, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

```
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
  source
  1..582
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="OFR09418"
    /tissue_type="mammary gland"
    /cell_line="OCUB-F"
    /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
    /note="mammary gland tumor"

ORGANISM
  Alignment Scores:
  Pred. No.: 7,596-18 Length: 582
  Score: 195.00 Matches: 38
  Percent Similarity: 100.00% Conservat: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP315806 (1-582)

QY 1 AsnLYsAlaLeuGluLeuLYsAsnGluInThrLeuArgAlaAspGluLeuProSer 20
264 AATAAGCCTTGGATTGAAGATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 323

QY 21 GluSerLYsGlnLYsAspTYrGluGluSerSerTrpAspSerGluSerLeuCys 38
324 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 377

Db
RESULT 8
BP312890
LOCUS
DEFINITION
  BP312890 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
  cDNA clone OFR01356, mRNA sequence.
ACCESSION
  BP312890.1 GI:52241865
VERSION
  BP312890.1
KEYWORDS
  EST.
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
  Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  Sequence comparison of human and mouse genes reveals a homologous
  block structure in the promoter regions
  Genome Res. 14 (9), 1711-1718 (2004)
  Contact: Yutaka Suzuki
  Department of Virology
  Institute of Medical Science, University of Tokyo
  4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
  Email: ysuzuki@ims.u-tokyo.ac.jp.
  Location/Qualifiers
    1..583
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="OFR01356"
      /tissue_type="mammary gland"
      /cell_line="OCUB-F"
      /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
      /note="mammary gland tumor"

REFERENCE
  1 (bases 1 to 583)
  Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
  Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  Sequence comparison of human and mouse genes reveals a homologous
  block structure in the promoter regions
  Genome Res. 14 (9), 1711-1718 (2004)
  Contact: Yutaka Suzuki
  Department of Virology
  Institute of Medical Science, University of Tokyo
  4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
  Email: ysuzuki@ims.u-tokyo.ac.jp.
  Location/Qualifiers
    1..583
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="OFR01356"
      /tissue_type="mammary gland"
      /cell_line="OCUB-F"
      /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
      /note="mammary gland tumor"

FEATURES
  source
  1..583
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="OFR01356"
    /tissue_type="mammary gland"
    /cell_line="OCUB-F"
    /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
    /note="mammary gland tumor"

ORGANISM
  Alignment Scores:
  Pred. No.: 7,61e-18 Length: 583
  Score: 195.00 Matches: 38
  Percent Similarity: 100.00% Conservat: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP312890 (1-583)

QY 1 AsnLYsAlaLeuGluLeuLYsAsnGluInThrLeuArgAlaAspGluLeuProSer 20
278 AATAAGCCTTGGATTGAAGATGAACAAACATTGAGAGCAGATGAGATCTCCATCA 337

QY 21 GluSerLYsGlnLYsAspTYrGluGluSerSerTrpAspSerGluSerLeuCys 38
338 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGGGATTCGAGAGTCTCTGT 391

Db
RESULT 9
BP313822
LOCUS
DEFINITION
  BP313822 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
  cDNA clone OFR04566, mRNA sequence.
ACCESSION
  BP313822
VERSION
  BP313822.1 GI:52242797
KEYWORDS
  EST.
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
  Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  Sequence comparison of human and mouse genes reveals a homologous
  block structure in the promoter regions
  Genome Res. 14 (9), 1711-1718 (2004)
  Contact: Yutaka Suzuki
  Department of Virology
  Institute of Medical Science, University of Tokyo
  4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
  Email: ysuzuki@ims.u-tokyo.ac.jp.
  Location/Qualifiers
    1..583
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="OFR04566"
      /tissue_type="mammary gland"
      /cell_line="OCUB-F"
      /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
      /note="mammary gland tumor"

REFERENCE
  1 (bases 1 to 583)
  Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
  Mizushima-Sugano,J., Nakai,K. and Sugano,S.
  Sequence comparison of human and mouse genes reveals a homologous
  block structure in the promoter regions
  Genome Res. 14 (9), 1711-1718 (2004)
  Contact: Yutaka Suzuki
  Department of Virology
  Institute of Medical Science, University of Tokyo
  4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
  Email: ysuzuki@ims.u-tokyo.ac.jp.
  Location/Qualifiers
    1..583
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="OFR04566"
      /tissue_type="mammary gland"
      /cell_line="OCUB-F"
      /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
      /note="mammary gland tumor"

FEATURES
  source
  1..583
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="OFR04566"
    /tissue_type="mammary gland"
    /cell_line="OCUB-F"
    /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
    /note="mammary gland tumor"

ORGANISM
  Alignment Scores:
  Pred. No.: 7,61e-18 Length: 583
  Score: 195.00 Matches: 38
  Percent Similarity: 100.00% Conservat: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP313822 (1-583)

QY 1 AsnLYsAlaLeuGluLeuLYsAsnGluInThrLeuArgAlaAspGluLeuProSer 20
|||||
```

Db 233 AATAAGCCTTGAATTGAAGATGAACAAACCTTGAGAGCAGATGAGATATCCCATCA 292
 QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrrpAspSerGluSerLeuCys 38
 |||||
 Db 293 GAATCCAAACAAAGAGCTATGAGAAAGTCTTGGGATTCGAGAGTCTCTGT 346

RESULT 10

BP314652
 LOCUS BP314652 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION cDNA clone OFR06752, mRNA sequence.

ACCESSION BP314652

VERSION BP314652.1 GI:52243627

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 584)

Mizushima-Sugano,J., Nakai,K. and Sugano,S.

TITLE Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

Genome Res. 14 (9), 1711-1718 (2004)

CONTACT: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES

source

1. .584

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="OFR06752"

/tissue_type="mammary gland"

/cell_line="OCUB-F"

/clone_lib="Sugano cDNA library, mammary gland OCUB-F"

/note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 7.63e-18 Length: 584
 Score: 195.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-489-079-27 (1-38) x BP314652 (1-584)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuLeuProSer 20
 |||||
 Db 278 AATAAGCCTTGAATTGAAGATGAACAAACCTTGAGAGCAGATGAGATATCCCATCA 337

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrrpAspSerGluSerLeuCys 38
 |||||
 Db 338 GAATCCAAACAAAGAGCTATGAGAAAGTCTTGGGATTCGAGAGTCTCTGT 391

RESULT 11

BP312910
 LOCUS BP312910 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION cDNA clone OFR01441, mRNA sequence.

ACCESSION BP312910

VERSION BP312910.1 GI:52241885

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 598)

Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,

TITLE

JOURNAL

COMMENT

CONTACT: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1. .598

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="OFR01441"

/tissue_type="mammary gland"

/cell_line="OCUB-F"

/clone_lib="Sugano cDNA library, mammary gland OCUB-F"

/note="mammary gland tumor"

ORIGIN

Alignment Scores:

Pred. No.: 7.85e-18 Length: 598
 Score: 195.00 Matches: 38
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0

US-09-489-079-27 (1-38) x BP312910 (1-598)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluLeuLeuProSer 20
 |||||
 Db 279 AATAAGCCTTGAATTGAAGATGAACAAACCTTGAGAGCAGATGAGATATCCCATCA 338

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrrpAspSerGluSerLeuCys 38
 |||||
 Db 339 GAATCCAAACAAAGAGCTATGAGAAAGTCTTGGGATTCGAGAGTCTCTGT 392

RESULT 12

BP746270
 LOCUS BP746270 443 bp mRNA linear EST 10-JAN-2001
 DEFINITION RC1-BT0254-071100-119-a03 BT0254 Homo sapiens cDNA, mRNA sequence.

ACCESSION BP746270

VERSION BP746270.1 GI:12072946

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 443)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jorgeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

PUBMED

COMMENT

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?ts=RC1&t2=RC1-BT0254-

071100-119-a03&t3=2000-11-07&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 9
 High quality sequence stop: 439.

FEATURES

Location/Qualifiers
 source
 1..443
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BT0254"
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 4,13e-16 Length: 443
 Score: 182.00 Matches: 37
 Percent Similarity: 97.37% Conservative: 0
 Best Local Similarity: 97.37% Mismatches: 1
 Query Match: 93.33% Indels: 0
 DB: 2 Gaps: 0

US-09-489-079-27 (1-38) x BF746270 (1-443)

QY 1 AenLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
 |||||
 Db 196 AATAAGCCTTGGAATTCGAGATGACAAACATTGAGAGCAGATGAGATCTCCATCA 255
 |||||
 QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
 |||||
 Db 256 GAATCCAAACAAAGGACTATGAGAAAGTTCTGGGGATTCCTGAGAGTCTCTGT 309
 |||||

RESULT 13

BF746340
 LOCUS 443 bp mRNA linear EST 10-JAN-2001
 DEFINITION RCL-BT0254-081100-119-a03 BT0254 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF746340
 VERSION BF746340.1 GI:12073016
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 443)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC1&t2=RC1-BT0254-081100-119-a03&t3=2000-11-08&t4=1)

Seq primer: puc 18 forward
 High quality sequence start: 9
 High quality sequence stop: 439.

FEATURES

Location/Qualifiers
 source
 1..443
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BT0254"
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 4,13e-16 Length: 443
 Score: 182.00 Matches: 37
 Percent Similarity: 97.37% Conservative: 0
 Best Local Similarity: 97.37% Mismatches: 1
 Query Match: 93.33% Indels: 0
 DB: 2 Gaps: 0

US-09-489-079-27 (1-38) x BF746340 (1-443)

QY 1 AenLysAlaLeuGluLeuLysAsnGluInThrLeuArgAlaAspGluLeuProSer 20
 |||||
 Db 196 AATAAGCCTTGGAATTCGAGATGACAAACATTGAGAGCAGATGAGATCTCCATCA 255
 |||||
 QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTrpAspSerGluSerLeuCys 38
 |||||
 Db 256 GAATCCAAACAAAGGACTATGAGAAAGTTCTGGGGATTCCTGAGAGTCTCTGT 309
 |||||

RESULT 14

BP313800
 LOCUS 583 bp mRNA linear EST 17-SEP-2004
 DEFINITION BP313800 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 ACCESSION BP313800
 VERSION BP313800.1 GI:52242775
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 583)
 Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: ysuzuki@ims.u-tokyo.ac.jp.

JOURNAL
 COMMENT Location/Qualifiers
 source
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OFR04504"
 /cell_type="mammary gland"
 /clone_lib="OCUB-F"
 /note="mammary gland tumor"

FEATURES

Location/Qualifiers
 source
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OFR04504"
 /cell_type="mammary gland"
 /clone_lib="OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:

```
Pred. No.: 1.12e-15 Length: 583
Score: 180.00 Matches: 37
Percent Similarity: 97.37% Conservative: 0
Best Local Similarity: 97.37% Mismatches: 1
Query Match: 92.31% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP313800 (1-583)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeuProSer 20
    |||||
DB 260 AATAAGCCCTTGGAAATTTGAGATGAAACAAACATTGAGAGCAGATGATCTCCCATCA 319
    |||||

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTyrAspSerGluSerLeuCys 38
    |||||
DB 320 GAATCCAAACAAAGGACTATGAAGAAAGTTCTTGAGATTCTGAGAGTCTCTGT 373
    |||||

RESULT 15
BP315089
LOCUS
DEFINITION BP315089 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
CDNA clone OFR07766, mRNA sequence.
ACCESSION BP315089
VERSION BP315089.1 GI:52244064
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR07766"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/notes="mammary gland tumor"

ORIGIN

Alignment Scores:
Pred. No.: 4.35e-14 Length: 582
Score: 169.00 Matches: 34
Percent Similarity: 92.11% Conservative: 1
Best Local Similarity: 89.47% Mismatches: 3
Query Match: 86.67% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-27 (1-38) x BP315089 (1-582)

QY 1 AsnLysAlaLeuGluLeuLysAsnGluGlnThrLeuArgAlaAspGluIleLeuProSer 20
    |||||
DB 322 AATAAGCCCTTGGAAATTTGAGATGAAACAAACATTGAGAGCAGATGATCTCCCATCA 381
    |||||

QY 21 GluSerLysGlnLysAspTyrGluGluSerSerTyrAspSerGluSerLeuCys 38
    |||||
DB 382 TAATCCAAACAAAGGACTATGAAGAAAGTTCTTGAGATTCTGAGAGTCTCTGT 435
    |||||

Search completed: August 2, 2005, 11:46:13
Job time : 1447.31 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:04:01 ; Search time 13.5071 Seconds
(without alignments)
1088.086 Million cell updates/sec

Title: US-09-489-079-27

Perfect score: 195

Sequence: 1 NKALEKNEQTLRADELILPSKQDYESSWDSLSL 38

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Genesecp1980s:*
- 2: Genesecp1990s:*
- 3: Genesecp2000s:*
- 4: Genesecp2001s:*
- 5: Genesecp2002s:*
- 6: Genesecp2003as:*
- 7: Genesecp2003bs:*
- 8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195	100.0	38	3 AAB07641	Aab07641 Amino aci
2	195	100.0	44	3 AAB07640	Aab07640 Amino aci
3	195	100.0	168	4 AAB63909	Aab63909 Human pro
4	195	100.0	169	4 AAB63906	Aab63906 Human pro
5	195	100.0	174	4 AAB63926	Aab63926 Human pro
6	195	100.0	185	4 AAB63933	Aab63933 Human pro
7	195	100.0	189	4 AAB63929	Aab63929 Human pro
8	195	100.0	195	4 AAB63918	Aab63918 Human pro
9	195	100.0	207	4 AAB63937	Aab63937 Human pro
10	195	100.0	220	4 AAB63917	Aab63917 Human pro
11	195	100.0	223	4 AAB63903	Aab63903 Human pro
12	195	100.0	225	4 AAB63901	Aab63901 Human pro
13	195	100.0	229	4 AAB63925	Aab63925 Human pro
14	195	100.0	241	7 ADB63983	Aab63983 Human bre
15	195	100.0	266	4 AAB63899	Aab63899 Human pro
16	195	100.0	398	3 AAB07638	Aab07638 Amino aci
17	195	100.0	445	4 AAB50249	Aab50249 Human bre
18	195	100.0	445	4 AAG65987	Aag65987 B726P spl
19	195	100.0	445	4 AAU33350	Aau33350 Human bre
20	195	100.0	445	5 ABG78917	Abg78917 Human bre
21	195	100.0	445	6 ABJ37740	Abj37740 Human tum
22	195	100.0	445	7 ADL93135	Adl93135 Human bre
23	195	100.0	445	8 ADE44425	Ade44425 Human bre
24	195	100.0	466	4 AAB50248	Aab50248 Human bre
25	195	100.0	466	4 AAG65986	Aag65986 B726P spl

26	195	100.0	466	4 AAU33349	Aau33349 Human bre
27	195	100.0	466	5 ABG78916	Abg78916 Human bre
28	195	100.0	466	6 ABJ37739	Abj37739 Human tum
29	195	100.0	466	7 ADL93134	Adl93134 Human bre
30	195	100.0	466	8 ADE44424	Ade44424 Human bre
31	195	100.0	512	4 AAB84701	Aab84701 Amino aci
32	195	100.0	650	4 AAB50263	Aab50263 Human bre
33	195	100.0	650	4 AAG65983	Aag65983 B726P spl
34	195	100.0	650	4 AAU33346	Aau33346 Human bre
35	195	100.0	650	5 ABG78913	Abg78913 Human bre
36	195	100.0	650	6 ABJ37736	Abj37736 Human tum
37	195	100.0	650	7 ADL93131	Adl93131 Human bre
38	195	100.0	650	8 ADE44421	Ade44421 Human bre
39	195	100.0	661	6 ABJ37782	Abj37782 Human tum
40	195	100.0	661	7 ADL93214	Adl93214 Human bre
41	195	100.0	743	4 AAU33358	Aau33358 Human bre
42	195	100.0	743	5 ABG78925	Abg78925 Human bre
43	195	100.0	743	6 ABJ37748	Abj37748 Human tum
44	195	100.0	743	7 ADL93156	Adl93156 Human bre
45	195	100.0	1002	4 AAU33351	Aau33351 Human bre

ALIGNMENTS

RESULT 1

AAB07641

ID AAB07641 standard; peptide; 38 AA.

AC AAB07641;

XX 07-NOV-2000 (first entry)

XX Amino acid sequence of an antigenic peptide from BS322 polypeptide.

XX BS322; breast tissue marker; breast disease; breast cancer.

XX Homo sapiens.

XX WO200043420-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-US001452.

XX 21-JAN-1999; 99US-00234716.

XX (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

XX Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;

XX Stroupe SD;

XX WPI; 2000-499217/44.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the detection and diagnosis of breast disease especially breast cancer.

XX Claim 23; Page 125; 126pp; English.

XX The present sequence represents an antigenic peptide derived from human BS322 polypeptide. The peptide is used to raise antibodies. BS322 is a breast tissue marker. The BS322 polynucleotides and polypeptides are used to detect and diagnose breast disease, e.g. breast cancer. The BS322 polynucleotides are useful as a source of probes and primers, and the BS322 polypeptides are useful as antigens

XX Sequence 38 AA;

XX Query Match 100.0%; Score 195; DB 3; Length 38;

XX Best Local Similarity 100.0%; Pred. No. 4.4e-19;

XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNQETLRADAILPSESQKQDYESSWDSLSLC 38
 XX |||||||
 DB 1 NKALELKNQETLRADAILPSESQKQDYESSWDSLSLC 38
 XX |||||||

RESULT 2

AAB07640
 ID AAB07640 standard; peptide; 44 AA.

XX AAB07640;

XX 07-NOV-2000 (first entry)

XX Amino acid sequence of an antigenic peptide from BS322 polypeptide.

DE BS322; breast tissue marker; breast disease; breast cancer.

KW Homo sapiens.

OS WO200043420-A1.

XX PN 27-JUL-2000.

XX PF 21-JAN-2000; 2000WO-US001452.

XX PR 21-JAN-1999; 99US-00234716.

XX PA (ABBO) ABBOTT LAB.

XX PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;

PI Stroupe SD;

XX WPI; 2000-499217/44.

DR BS322-specific polynucleotides, polypeptides and antibodies, used in the
 XX detection and diagnosis of breast disease especially breast cancer.

PS Claim 23; Page 125; 126pp; English.

XX The present sequence represents an antigenic peptide derived from human
 CC BS322 polypeptide. The peptide is used to raise antibodies. BS322 is a
 CC breast tissue marker. The BS322 polynucleotides and polypeptides are used
 CC to detect and diagnose breast disease, e.g. breast cancer. The BS322
 CC polynucleotides are useful as a source of probes and primers, and the
 CC BS322 polypeptides are useful as antigens

XX Sequence 44 AA;

Query Match 100.0%; Score 195; DB 3; Length 44;
 Best Local Similarity 100.0%; Pred. No. 5.3e-19; Mismatches 0; Indels 0; Gaps 0;
 Matches 38; Conservative 0;

QY 1 NKALELKNQETLRADAILPSESQKQDYESSWDSLSLC 38
 DB 7 NKALELKNQETLRADAILPSESQKQDYESSWDSLSLC 44
 XX |||||||

RESULT 3

AAB63909
 ID AAB63909 standard; protein; 168 AA.

XX AAB63909;

XX 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1271.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytostatic; cancer vaccine.

XX Homo sapiens.

OS WPI; 2001-025274/03.

XX DR

PN WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014749.

XX 28-MAY-1999; 99US-0136526P.

PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

PI WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.

PT cancer.

XX Example 1; Page 772; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

CC represent human breast, gastric and prostate CAAP protein sequence

CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterised by expression of an abnormal amount of a protein,

CC e.g. cancer

XX SQ Sequence 168 AA;

Query Match 100.0%; Score 195; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 2.9e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNQETLRADAILPSESQKQDYESSWDSLSLC 38
 DB 113 NKALELKNQETLRADAILPSESQKQDYESSWDSLSLC 150
 XX |||||||

RESULT 4

AAB63906
 ID AAB63906 standard; protein; 169 AA.

XX AAB63906;

XX 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1268.

DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytostatic; cancer vaccine.

XX Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014749.

XX 28-MAY-1999; 99US-0136526P.

PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX DR

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 XX Example 1; Page 770-771; 799pp; English.
 PS
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterized by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 169 AA;

Query Match 100.0%; Score 195; DB 4; Length 169;
 Best Local Similarity 100.0%; Pred. No. 2.9e-18; Indels 0; Gaps 0;
 Matches 38; Conservative 0; Mismatches 0;

QY 1 NKALELKNQETLRADILPESKQDYESSWDSLSL 38
 DB 97 NKALELKNQETLRADILPESKQDYESSWDSLSL 134

RESULT 5
 AAB63926
 ID AAB63926 standard; protein; 174 AA.
 XX
 XX AAB63926;
 XX
 XX 26-MAR-2001 (first entry)
 XX
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1288.
 DE
 XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 XX Homo sapiens.
 OS
 XX WO200073801-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US014749.
 PP
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Obata Y;
 PI
 XX WPI; 2001-025274/03.
 DR
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.

XX Example 1; Page 779; 799pp; English.
 PS
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterized by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 174 AA;

Query Match 100.0%; Score 195; DB 4; Length 174;
 Best Local Similarity 100.0%; Pred. No. 3e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNQETLRADILPESKQDYESSWDSLSL 38
 DB 25 NKALELKNQETLRADILPESKQDYESSWDSLSL 62

RESULT 6
 AAB63933
 ID AAB63933 standard; protein; 185 AA.

XX
 XX AAB63933;
 XX
 XX 26-MAR-2001 (first entry)
 DT
 XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1295.

DE
 XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 XX Homo sapiens.
 OS

XX WO200073801-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US014749.
 PP
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Obata Y;
 PI
 XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 XX Example 1; Page 782; 799pp; English.
 PS
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterized by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 185 AA;

Query Match 100.0%; Score 195; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 3.2e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNQETLRADILPESKQDYESSWDSLSL 38
 DB 96 NKALELKNQETLRADILPESKQDYESSWDSLSL 133

```

RESULT 7
AAB63929
ID AAB63929 standard; protein; 189 AA.
XX AC AAB63929;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1291.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
PS Example 1; Page 780; 799pp; English.
XX
CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX
SQ Sequence 189 AA;
XX
PS Example 1; Page 780; 799pp; English.
XX
CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX
SQ Sequence 189 AA;
XX
Query Match 100.0%; Score 195; DB 4; Length 189;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 NKALELKNQTLRADEILPSESQKDYESSWDSLSL 38
DB 118 NKALELKNQTLRADEILPSESQKDYESSWDSLSL 155
XX
RESULT 8
AAB63918
ID AAB63918 standard; protein; 195 AA.
XX AC AAB63918;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1280.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX

```

```

XX OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
PS Example 1; Page 775-776; 799pp; English.
XX
CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer
XX
SQ Sequence 195 AA;
XX
Query Match 100.0%; Score 195; DB 4; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.5e-18;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 NKALELKNQTLRADEILPSESQKDYESSWDSLSL 38
DB 84 NKALELKNQTLRADEILPSESQKDYESSWDSLSL 121
XX
RESULT 9
AAB63937
ID AAB63937 standard; protein; 207 AA.
XX AC AAB63937;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1299.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX

```

PI Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 XX Example 1; Page 783-784; 799pp; English.
 XX
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 207 AA;
 SQ
 Query Match 100.0%; Score 195; DB 4; Length 207;
 Best Local Similarity 100.0%; Pred. No. 3.7e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NKALELKNQTLRADEILPESKQDYESSWDSLSL 38
 DB 17 NKALELKNQTLRADEILPESKQDYESSWDSLSL 54
 XX
 RESULT 10
 AAB63917
 ID AAB63917 standard; protein; 220 AA.
 XX
 AC AAB63917;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1279.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 DR WPI; 2001-025274/03.
 XX
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 XX Example 1; Page 775; 799pp; English.
 XX
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 220 AA;
 SQ
 Query Match 100.0%; Score 195; DB 4; Length 220;
 Best Local Similarity 100.0%; Pred. No. 4e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 NKALELKNQTLRADEILPESKQDYESSWDSLSL 38
 DB 14 NKALELKNQTLRADEILPESKQDYESSWDSLSL 51
 XX
 RESULT 11
 AAB63903
 ID AAB63903 standard; protein; 223 AA.
 XX
 AC AAB63903;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1265.
 XX
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 DR WPI; 2001-025274/03.
 XX
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX
 XX Example 1; Page 769; 799pp; English.
 XX
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX
 XX Sequence 223 AA;
 SQ
 Query Match 100.0%; Score 195; DB 4; Length 223;
 Best Local Similarity 100.0%; Pred. No. 4.1e-18;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 38
 |||||
 Db 104 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 141

RESULT 12
 AAB63901
 ID AAB63901 standard; protein; 225 AA.
 XX
 AC AAB63901;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1263.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 WPI; 2001-025274/03.
 XX

XX Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.

PT cancer.
 XX

PS Example 1; Page 768; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and

CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

CC represent human breast, gastric and prostate CAAP protein sequence

CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX

SQ Sequence 225 AA;

Query Match 100.0%; Score 195; DB 4; Length 225;

Best Local Similarity 100.0%; Pred. No. 4.1e-18;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 38
 |||||
 Db 22 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 59

RESULT 13
 AAB63925
 ID AAB63925 standard; protein; 229 AA.
 XX
 AC AAB63925;
 XX
 DT 26-MAR-2001 (first entry)
 XX
 DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1287.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 WPI; 2001-025274/03.
 XX

XX Nucleic acids encoding breast, gastric and prostate cancer associated

PT antigen precursors, useful for diagnosing and treating a condition

PT characterized by expression of an abnormal amount of a protein, e.g.

PT cancer.
 XX

PS Example 1; Page 779; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and

CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

CC represent human breast, gastric and prostate CAAP protein sequence

CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX

SQ Sequence 229 AA;

Query Match 100.0%; Score 195; DB 4; Length 229;

Best Local Similarity 100.0%; Pred. No. 4.2e-18;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 38
 |||||
 Db 25 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 62

RESULT 14
 ADB83983
 ID ADB83983 standard; protein; 241 AA.
 XX
 AC ADB83983;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human breast cancer diagnostic marker protein Incyte 411152.
 XX
 KW cancer; neurodegenerative disorder; human; breast cancer;
 KW diagnostic marker.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by GAG"
 FT Misc-difference 75 /note= "Encoded by GAT"
 FT Misc-difference 76 /note= "Encoded by AAT"
 FT Misc-difference 79

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200073801-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 26-MAY-2000; 2000WO-US014749.
 XX
 PR 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX
 PA (LUDW-) LUDWIG INST CANCER RES.
 XX
 PI Obata Y;
 XX
 WPI; 2001-025274/03.
 XX

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX

PS Example 1; Page 778-779; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014

CC represent nucleotide sequences encoding human breast, gastric and

CC prostate cancer associated antigen precursors (CAAP) respectively.

CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970

CC represent human breast, gastric and prostate CAAP protein sequence

CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic

CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX

SQ Sequence 229 AA;

Query Match 100.0%; Score 195; DB 4; Length 229;

Best Local Similarity 100.0%; Pred. No. 4.2e-18;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 38
 |||||
 Db 25 NKALELKNEQTLRADEILPSESKQDYESSWDSLESLC 62

RESULT 14
 ADB83983
 ID ADB83983 standard; protein; 241 AA.
 XX
 AC ADB83983;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human breast cancer diagnostic marker protein Incyte 411152.
 XX
 KW cancer; neurodegenerative disorder; human; breast cancer;
 KW diagnostic marker.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by GAG"
 FT Misc-difference 75 /note= "Encoded by GAT"
 FT Misc-difference 76 /note= "Encoded by AAT"
 FT Misc-difference 79

```

FT      /note= "Encoded by TTT"
FT      Misc-difference 83
FT      /note= "Encoded by CCC"
FT      Misc-difference 85
FT      /note= "Encoded by AGA"
XX
XX      US2003104418-A1.
XX
XX      05-JUN-2003.
XX
XX      25-APR-2002; 2002US-001373757.
XX
XX      27-APR-2001; 2001US-0287153P.
XX      (ZHAN/) ZHANG C.
XX      (MAHI/) MAHINI B.
XX      (WALK/) WALKER M G.
XX
XX      Zhang C, Mahini B, Walker MG;
XX
XX      WPI; 2003-687833/65.
XX      N-PSDB; ADB83986.
XX
XX      New combination of polynucleotides, useful for preparing a composition
XX      for diagnosing or treating cancer or neurodegenerative disorders.
XX
XX      Example 11; Page 20-21; 22pp; English.
XX
XX      The invention related to a combination of polynucleotides. The
XX      combination of polynucleotides is useful for preparing a composition for
XX      diagnosing or treating cancer or neurodegenerative disorders. The present
XX      sequence represents the amino acid sequence of the human breast cancer
XX      diagnostic marker Incyte 411152.
XX
XX      Sequence 241 AA;
SQ
      Query Match      100.0%; Score 195; DB 7; Length 241;
      Best Local Similarity 100.0%; Pred. No. 4.5e-18;
      Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NKALELKNQTLRADEILPSESQKQDYESSWDSLSLC 38
DB      |||||||
      7 NKALELKNQTLRADEILPSESQKQDYESSWDSLSLC 44

RESULT 15
AAB63899
XX      AAB63899 standard; protein; 266 AA.
XX
XX      AAB63899;
XX
XX      26-MAR-2001 (first entry)
XX
XX      Human prostate cancer associated antigen protein sequence SEQ ID NO:1261.
XX
XX      Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
XX      cancer associated antigen; cytostatic; cancer vaccine.
XX
XX      Homo sapiens.
XX
XX      WO200073801-A2.
XX
XX      07-DEC-2000.
XX
XX      26-MAY-2000; 2000WO-US014749.
XX
XX      28-MAY-1999; 99US-0136526P.
XX      10-SEP-1999; 99US-0153454P.
XX
XX      (LUDW-) LUDWIG INST CANCER RES.
XX
XX      Obata Y;
XX

```

```

DR      WPI; 2001-025274/03.
XX
XX      Nucleic acids encoding breast, gastric and prostate cancer associated
XX      antigen precursors, useful for diagnosing and treating a condition
XX      characterized by expression of an abnormal amount of a protein, e.g.
XX      cancer.
XX
XX      Example 1; Page 767; 799pp; English.
XX
XX      AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
XX      represent nucleotide sequences encoding human breast, gastric and
XX      prostate cancer associated antigen precursors (CAAP) respectively.
XX      AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
XX      represent human breast, gastric and prostate CAAP protein sequence
XX      respectively. CAAPs have cytostatic activity and can be used in the
XX      production of cancer vaccines. The human CAAP proteins, peptides, nucleic
XX      acids or anti-CAAP antibodies are useful for diagnosing and treating a
XX      condition characterised by expression of an abnormal amount of a protein,
XX      e.g. cancer
XX
XX      Sequence 266 AA;
SQ
      Query Match      100.0%; Score 195; DB 4; Length 266;
      Best Local Similarity 100.0%; Pred. No. 5.1e-18;
      Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 NKALELKNQTLRADEILPSESQKQDYESSWDSLSLC 38
DB      |||||||
      28 NKALELKNQTLRADEILPSESQKQDYESSWDSLSLC 65

Search completed: August 1, 2005, 22:18:04
Job time : 13.5071 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:05:16 ; Search time 10.6256 Seconds
(without alignments)
1831.334 Million cell updates/sec

Title: US-09-489-079-27
Perfect score: 195
Sequence: 1 NKALELKNQTLRADEILPSESKQDYESSWDSLSLC 38

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues
Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	100.0	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	153	78.5	1011	2 Q9BXX2	Q9bxx2 homo sapien
3	65	33.3	823	2 Q9H0H6	Q9h0h6 homo sapien
4	62	31.8	223	2 Q9QZM8	Q9qzm8 rattus norv
5	61	31.3	226	2 Q9D5A2	Q9d5a2 mus musculu
6	61	31.3	234	2 Q802P0	Q80zp0 mus musculu
7	61	31.3	851	2 Q81IE2	Q81ie2 plasmodium
8	60.5	31.0	899	2 Q9SG75	Q9sg75 arabidopsis
9	60.5	31.0	917	2 Q9CAF6	Q9caf6 arabidopsis
10	60	30.8	391	2 Q8KRC9	Q8krc9 myxococcus
11	59	30.3	380	2 Q6LXV6	Q6lxv6 methanococc
12	58.5	30.0	678	2 Q17399	Q17399 caenorhabdi
13	58.5	30.0	1319	2 Q9HCD3	Q9hcd3 homo sapien
14	58	29.7	548	2 Q7XPN5	Q7xpn5 oryza sativ
15	58	29.7	1066	2 Q7XE90	Q7xe90 oryza sativ
16	58	29.7	1654	2 Q7RHE8	Q7rhe8 plasmodium
17	57	29.2	568	2 Q7S711	Q7s711 neurospora
18	57	29.2	1806	2 Q6LFB1	Q6lfb1 plasmodium
19	56.5	29.0	707	2 Q913P9	Q913p9 pseudomonas
20	56.5	29.0	1487	2 Q7RBD9	Q7rbd9 plasmodium
21	56	28.7	259	2 Q9CWH4	Q9cwh4 mus musculu
22	56	28.7	294	2 Q6UGS2	Q6ugg2 enterobacte
23	56	28.7	697	2 Q9NXX9	Q9nyj9 homo sapien
24	56	28.7	703	2 Q91Y19	Q91y19 mus musculu
25	56	28.7	719	2 Q6PIQ6	Q6piq6 homo sapien
26	56	28.7	752	1 HPR1 YEAST	P17629 saccharomyc
27	56	28.7	1095	2 Q9HCG6	Q9hcg6 homo sapien
28	56	28.7	1720	2 Q6E7C8	Q6e7c8 oikopleura
29	56	28.7	1948	2 Q6LFI3	Q6lfi3 plasmodium
30	56	28.7	2063	2 Q80TB4	Q80tb4 mus musculu
31	55.5	28.5	125	2 Q973D7	Q973d7 sulfolobus

32	55.5	28.5	237	1 YFHG_ECOLI	P37328 escherichia
33	55.5	28.5	332	1 SP2B_BACSU	P37375 bacillus su
34	55.5	28.5	428	2 Q7MA44	Q7ma44 wolinnella s
35	55.5	28.5	506	2 Q7N714	Q7n714 photorhabdu
36	55.5	28.5	558	2 O56966	O56966 lymphocytic
37	55.5	28.5	725	2 Q7RHR7	Q7rhr7 plasmodium
38	55.5	28.5	746	2 Q7MA46	Q7ma46 wolinnella s
39	55.5	28.5	860	2 Q9M9Z7	Q9m9z7 arabidopsis
40	55.5	28.5	1016	2 Q8LPG5	Q8lpg5 arabidopsis
41	55.5	28.5	1016	2 Q8LPR9	Q8lpr9 arabidopsis
42	55.5	28.5	2276	2 Q79LN3	Q79ln3 staphylococ
43	55	28.2	134	2 Q7PSL6	Q7psl6 fusobacteri
44	55	28.2	185	2 Q64PG9	Q64pg9 bacteroides
45	55	28.2	311	2 Q6FTF4	Q6ftf4 candida gla

ALIGNMENTS

RESULT 1
Q9BXX3 PRELIMINARY; PRT; 1341 AA.
ID Q9BXX3
AC Q9BXX3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Breast cancer antigen NY-BR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library."
RL Cancer Res. 61:2055-2061 (2001).
DR EMBL; AF269087; AAK27325.1; -;
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0003700; F:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
DR ANK repeat.
SQ SEQUENCE 1341 AA; 152776 MW; 33E53DDE6FD3A58B CRC64;
Query Match 100.0%; Score 195; DB 2; Length 1341;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NKALELKNQTLRADEILPSESKQDYESSWDSLSLC 38
Db 717 NKALELKNQTLRADEILPSESKQDYESSWDSLSLC 754
RESULT 2
Q9BXX2 PRELIMINARY; PRT; 1011 AA.
ID Q9BXX2
AC Q9BXX2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
RA Jaeger E., Knuth A., Old L.J., Chen Y.T.,
RT "Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
RL Cancer Res. 61:2055-2061 (2001).
DR EMBL; AF269088; AAK27326.1; -.
DR HSSP; O75832; IUOH.
DR InterPro; IPR002110; ANK.
DR SMART; PF00023; ANK; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 1.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat.
FT NON TER 1011 1011
SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DE438 CRC64;

Query Match 78.5%; Score 153; DB 2; Length 1011;
Best Local Similarity 78.9%; Pred. No. 1.2e-09;
Matches 30; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 NKALELKNQTLRADEILPSESQKQDYESSWDSLSL 38
DB 391 NKAFELKNQTLRAQMFPSQKQDYESSWDSLSPC 428

RESULT 3
Q9H0H6 PRELIMINARY; PRT; 823 AA.
ID Q9H0H6;
AC Q9H0H6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp434A171.
GN Name=DKFZp434A171;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RG The German cDNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AL136793; CAB66727.1; -.
DR HSSP; Q60778; I0Y3.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ANK; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
KW ANK repeat; Hypothetical protein.
SQ SEQUENCE 823 AA; 93968 MW; 9D2DD5F52C379B2 CRC64;

Query Match 33.3%; Score 65; DB 2; Length 823;
Best Local Similarity 50.0%; Pred. No. 38;
Matches 15; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 3 ALELKNQTLRADEILP--SESKQDYESS 30
DB 349 AVQRKNVQTLRAEQALPVANSEEQERHERS 378

RESULT 4
Q9QZW8 PRELIMINARY; PRT; 223 AA.
ID Q9QZW8;
AC Q9QZW8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DnaJ-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wistar; TISSUE=Testis;
RA Wang L., Miao S., Yang J., Zhang X., Zhao M., Li M.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF154849; AAD53061.1; -.
DR HSSP; P25685; 1HDIJ.
DR InterPro; IPR001623; DnaJ_N.
DR Pfam; PF00226; DnaJ; 1.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
SQ SEQUENCE 223 AA; 25756 MW; EBB4ADF3CC92D3F CRC64;

Query Match 31.8%; Score 62; DB 2; Length 223;
Best Local Similarity 41.9%; Pred. No. 21;
Matches 13; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 2 KALELKNQTLRADEILPSESQKQDYESSW 32
DB 40 EAABEKFKQVAEYQILLSDAKKRKYDRSRW 70

RESULT 5
Q9DS52 PRELIMINARY; PRT; 226 AA.
ID Q9DS52;
AC Q9DS52;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Mus musculus adult male testis cDNA, RIKEN full-length enriched
DE library, clone:4930483N21 product:similar to DNAJ-LIKE PROTEIN.
GN Name=4930503B20Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Testis;
RA MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Testis;
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Testis;
RA The FANTOM Consortium,
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Testis;
RL MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

```



```
Best Local Similarity 37.5%; Pred. No. 1.2e+02;
Matches 12; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 2 KALELKNQTLRADEILPSESKQDYESSWDS 33
Db 10 KGLETKNETIIEKENKEVEEKEFESEYN 41

RESULT 8
Q9SG75 PRELIMINARY; PRT; 899 AA.
AC Q9SG75;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative DNA gyrase A subunit.
GN Name=rvM13.23;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.,
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011708; AAF19580.1; -.
DR HSSP; P09097; IAB4.
DR GO; GO:0005694; C:chromosome; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR GO; GO:0006268; P:DNA unwinding; IEA.
DR InterPro; IPR005743; DNA_gyrA.
DR InterPro; IPR006691; DNA_gyraseA_C.
DR InterPro; IPR002205; DNA_topoisoIV.
DR Pfam; PF03989; DNA_gyraseA_C; 5.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR PRODOM; PD000742; DNA_topoisoIV; 1.
DR SMART; SM00434; TOP4c; 1.
DR TIGRFAMs; TIGR01063; gyra; 1.
SQ SEQUENCE 899 AA; 99547 MW; 1AC44F32B36E8333 CRC64;

Query Match 31.0%; Score 60.5; DB 2; Length 899;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 14; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

QY 3 ALELKNQTLRADEILPSESKQDYESSWDS 34
Db 767 AMLKNEKDIASMDIIPA-SLRKWEKSEDA 797

RESULT 9
Q9CAF6 PRELIMINARY; PRT; 917 AA.
AC Q9CAF6;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Putative DNA gyrase subunit A; 1114-7603.
GN Name=F13M14.2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxID=3702;
RN
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Matti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;

Query Match 31.0%; Score 60.5; DB 2; Length 899;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 14; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

QY 3 ALELKNQTLRADEILPSESKQDYESSWDS 34
Db 767 AMLKNEKDIASMDIIPA-SLRKWEKSEDA 797

RESULT 10
Q8KRC9 PRELIMINARY; PRT; 391 AA.
AC Q8KRC9;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE DnaJ.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophactereae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=DZF1;
RA Ueki T., Inouye S.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts as a co-chaperone. Stimulates, jointly with grpe,
CC the ATPase activity of dnaJ (By similarity).
CC -!- COFACTOR: Binds 2 zinc ions per monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the dnaJ family.
DR EMBL; AY033943; AAK59395.1; -.
DR HSSP; P25685; 1HDJ.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR GO; GO:0006986; P:response to unfolded protein; IEA.
DR InterPro; IPR002939; DnaJ_C.
DR InterPro; IPR001305; DnaJ_CXKXGKXG.
DR InterPro; IPR001623; DnaJ_N.
DR InterPro; IPR008971; HSP40_DnaJ_dep.
DR InterPro; IPR003095; Hsp_DnaJ.
DR Pfam; PF00226; DnaJ; 1.
DR Pfam; PF01556; DnaJ_C; 1.
DR Pfam; PF00684; DnaJ_CXKXGKXG; 1.
DR PRINTS; PR00625; DNAPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS00076; DnaJ_2; 1.
KW Chaperone; DNA replication; Heat shock; Metal-binding; Repeat; Zinc.
SQ SEQUENCE 391 AA; 40914 MW; 195A923B8618BFCE CRC64;
```

```
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011560; AAG51377.1; -.
DR HSSP; P09097; IAB4.
DR GO; GO:0005694; C:chromosome; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR GO; GO:0006268; P:DNA unwinding; IEA.
DR InterPro; IPR005743; DNA_gyrA.
DR InterPro; IPR006691; DNA_gyraseA_C.
DR InterPro; IPR002205; DNA_topoisoIV.
DR Pfam; PF03989; DNA_gyraseA_C; 5.
DR Pfam; PF00521; DNA_topoisoIV; 1.
DR PRODOM; PD000742; DNA_topoisoIV; 1.
DR SMART; SM00434; TOP4c; 1.
DR TIGRFAMs; TIGR01063; gyra; 1.
SQ SEQUENCE 917 AA; 101412 MW; 7568C4004D524976 CRC64;
```

```
Query Match 31.0%; Score 60.5; DB 2; Length 917;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 14; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

QY 3 ALELKNQTLRADEILPSESKQDYESSWDS 34
Db 785 AMLKNEKDIASMDIIPA-SLRKWEKSEDA 815
```

RP SEQUENCE FROM N.A.

DR FROST, F350835,
DR FROST: PS50835:

DR PROSITE, PS50835: IG LIKE: 1.
DR PROSITE, PS50835: LNS, 1.

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:23:28 ; Search time 87.0391 Seconds
(without alignments)
3196.585 Million cell updates/sec

Title: US-09-489-079-28
Perfect score: 246
Sequence: 1 KDGLKANGCMKVSIPTRKAL.....FKAGKFCNFNFTLRLILKY 47

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool_p/US09489079/runat_01082005_141249_13290/app_query.fasta_1.1635
-DB=N Geneseq 16Dec04 -QFT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN_1_1418 @runat_01082005_141249_13290 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOPEXT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2002bs: *
8: Geneseqn2003as: *
9: Geneseqn2003bs: *
10: Geneseqn2003cs: *
11: Geneseqn2003ds: *
12: Geneseqn2004as: *
13: Geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	152	61.8	403	5	ABV04090 Human pro
2	152	61.8	404	5	ABV13259 Human pro
3	152	61.8	415	11	ACN90082 Breast ca
4	152	61.8	419	10	ABT22964 Breast ca
5	152	61.8	434	4	AAL18710 Human bre

6	152	61.8	444	5	ABV28487
7	152	61.8	444	5	ABV34378
8	152	61.8	444	5	ABV22662
9	152	61.8	444	5	ABV43239
10	152	61.8	454	4	AAL10911
11	152	61.8	466	5	ABV13085
12	152	61.8	470	11	ACN81108
13	152	61.8	473	5	ABV43067
14	152	61.8	473	5	ABV34206
15	152	61.8	508	5	ABV03916
16	152	61.8	555	5	ABV41632
17	152	61.8	555	12	ADO40108
18	152	61.8	600	11	ACN92774
19	152	61.8	635	11	ACN87508
20	152	61.8	636	4	AAF22981
21	152	61.8	664	4	AAF22980
22	152	61.8	673	11	ACN86653
23	152	61.8	708	4	AAF22972
24	152	61.8	713	4	AAF22974
25	152	61.8	741	4	AAF22991
26	152	61.8	763	4	AAL18922
27	152	61.8	766	4	AAF22986
28	152	61.8	771	4	AAL20493
29	152	61.8	771	4	AAF22989
30	152	61.8	776	4	AAF22969
31	152	61.8	779	5	AAD06847
32	152	61.8	814	4	AAF22967
33	152	61.8	833	4	AAF22987
34	152	61.8	904	4	AAF22965
35	152	61.8	1158	4	AAF22985
36	152	61.8	1337	4	AAF17979
37	152	61.8	1337	4	AAL167222
38	152	61.8	1337	4	AAS47409
39	152	61.8	1337	6	ABS64010
40	152	61.8	1337	10	ABT33222
41	152	61.8	1337	11	ADL93129
42	152	61.8	1337	12	ADE44419
43	152	61.8	1448	5	ABV22558
44	152	61.8	1448	5	ABV28376
45	152	61.8	1448	5	ABV27350

ALIGNMENTS

RESULT 1
ABV04090
ID ABV04090 standard; cDNA; 403 BP.
XX
AC ABV04090;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 4081.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

```

XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 717; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 403 BP; 138 A; 79 C; 98 G; 87 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 3.25e-15 Length: 403
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV04090 (1-403)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
DB 173 AAAGATGGTCTCTCGAAGCTAACTGCGGATGAAGTTTCTATTCCAACTAAAGCCTTA 232

QY 21 GluLeuMetAspMetGlnThrPhelLysAla 30
DB 233 GAATTGATGGACATGCACAACTTTCAAAGCA 262

RESULT 2
ABV13259
ID ABV13259 standard; cDNA; 404 BP.
XX
XX AC ABV13259;
XX
XX DT 13-SEP-2002 (first entry)
XX
XX DE Human prostate expression marker cDNA 13250.
XX
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200160860-A2.
XX
XX PD 23-AUG-2001.
XX
XX PF 20-FEB-2001; 2001WO-US005171.
XX
XX PR 17-FEB-2000; 2000US-0183319P.
XX
XX PR 16-MAR-2000; 2000US-0189862P.
XX
XX PR 25-MAY-2000; 2000US-0207454P.
XX
XX PR 09-JUN-2000; 2000US-0211314P.
XX
XX PR 18-JUL-2000; 2000US-0219007P.
XX
XX PR 13-DEC-2000; 2000US-0255281P.
XX
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

```

```

XX Schlegel R, Endege WO, Monahan JE;
PI
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 2195; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 404 BP; 138 A; 79 C; 100 G; 87 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.26e-15 Length: 404
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV13259 (1-404)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
DB 173 AAAGATGGTCTCTCGAAGCTAACTGCGGATGAAGTTTCTATTCCAACTAAAGCCTTA 232

QY 21 GluLeuMetAspMetGlnThrPhelLysAla 30
DB 233 GAATTGATGGACATGCACAACTTTCAAAGCA 262

RESULT 3
ACN90082
ID ACN90082 standard; DNA; 415 BP.
XX
XX AC ACN90082;
XX
XX DT 02-DEC-2004 (first entry)
XX
XX DE Breast cancer related marker, seq id 11232.
XX
XX KW Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.
XX
XX OS Homo sapiens.
XX
XX PN US2003099974-A1.
XX
XX PD 29-MAY-2003.
XX
XX PF 18-JUL-2002; 2002US-00198846.
XX
XX PR 18-JUL-2001; 2001US-0306220P.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2003-787014/74.
XX
XX Novel isolated polypeptide associated with breast cancer, useful for

```

PT detecting presence of polypeptide in sample, as a marker for breast
PT cancer.
XX
PS Disclosure; SEQ ID NO 11232; 36pp; English.
XX
CC The invention relates to an isolated polypeptide (I) associated with a
CC breast cancer which is encoded by a nucleic acid molecule comprising a
CC nucleotide sequence (S1). Further disclosed is an antibody that binds to
CC the polypeptide of the invention. The activity of the polypeptide of the
CC invention may be described as cytostatic. The antibody is useful for
CC detecting the presence of (I) in a sample. Nucleic acid molecules of the
CC invention are useful in the detection of breast tumours. (I) is useful as
CC a marker for breast cancer and in breast cancer therapy. Sequences given
CC in records ACN78851-ACN92934 represent nucleic acid markers associated
CC with breast cancer. Note: The sequence listing does not form part of the
CC specification but may be obtained in electronic format from the USPTO web
CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974
XX
SQ Sequence 415 BP; 139 A; 73 C; 83 G; 119 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 3 37e-15 Length: 415
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 11 Gaps: 0

US-09-489-079-28 (1-47) x ACN90082 (1-415)

QY 1 LysAspGlyLeuLeuLysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db 299 AAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTA 358

QY 21 GluLeuMetAspMetGlnThrPhelLysAla 30
Db 359 GAATTGATGACATGCANAACTTTTCAAAGCA 388

RESULT 4
ABT22964
ID ABT22964 standard; DNA; 419 BP.
XX
AC ABT22964;
XX
DT 16-APR-2003 (first entry)
XX
DE Breast cancer marker gene SEQ ID No 1337.
XX
KW Cytostatic; vaccine; breast cancer marker gene; breast mass; immunogen;
KW chemotherapy; tumour burden; bait protein; two-hybrid; three-hybrid;
KW surrogate marker gene; pharmacodynamic marker gene; transgenic animal;
KW human; ds.
XX
OS Homo sapiens.
XX
PN WO200285298-A2.
XX
PD 31-OCT-2002.
XX
PF 19-APR-2002; 2002WO-US012612.
XX
PR 20-APR-2001; 2001US-0285163P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Lillie J, Palermo A, Wang Y, Steinmann K, Elias J, Mertens M;
XX
XX WPI; 2003-093053/08.
XX
PT Novel isolated polypeptide encoded by breast cancer marker gene, useful
PT for diagnosing, staging, monitoring, prognosing and treating diseases
PT associated with breast cancer.
XX

PS Disclosure; Page 271; 725pp; English.
XX
CC The invention relates to an isolated polypeptide encoded by a breast
CC cancer marker gene comprising any of 1417 21-805 nucleotide sequences,
CC given in the specification. The methods of the invention are useful for
CC diagnosing patients having an identified breast mass or symptoms
CC associated with breast cancer, to diagnose breast cancer or its
CC precursors, and for monitoring the efficacy of treatment of a breast
CC cancer patient (e.g. efficacy of chemotherapy). The methods are also
CC useful for evaluating a patient before, after or during therapy, to
CC evaluate the reduction in a tumour burden. The breast cancer marker gene
CC proteins are useful as immunogens for raising antibodies, by immunising a
CC mammal with a breast cancer marker protein. The marker proteins are
CC useful as bait proteins in a two-hybrid or three-hybrid assay, to
CC identify other proteins which bind to or interact with the marker
CC proteins. The breast cancer marker genes are useful as surrogate marker
CC genes for one or more disorders, disease states or conditions leading to
CC disease states, in particular, breast cancers. The breast cancer marker
CC genes are useful as pharmacodynamic marker genes. An antibody which
CC selectively binds to a protein of a breast cancer marker gene is useful
CC for treating cancers, particularly breast cancers. The host cell of the
CC invention is useful for producing non-human transgenic animals. This
CC polynucleotide sequence represents one of the breast cancer marker genes
CC of the invention
XX
SQ Sequence 419 BP; 146 A; 84 C; 94 G; 95 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3 42e-15 Length: 419
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 10 Gaps: 0

US-09-489-079-28 (1-47) x ABT22964 (1-419)

QY 1 LysAspGlyLeuLeuLysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db 196 AAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTA 255

QY 21 GluLeuMetAspMetGlnThrPhelLysAla 30
Db 256 GAATTGATGACATGCANAACTTTTCAAAGCA 285

RESULT 5
AAL18710
ID AAL18710 standard; cDNA; 434 BP.
XX
AC AAL18710;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 11167.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
XX WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer.

XX
XX Claim 1; Page 1991; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity

SQ Sequence 434 BP; 141 A; 81 C; 92 G; 120 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.58e-15 Length: 434
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-28 (1-47) x AAL18710 (1-434)

Qy 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrIysAlaLeu 20
Db 318 AAGATGGTCTCTCGAGGCTAACTGCGGAATGAAGTTCTATTCCAAACGCTTA 377

Qy 21 GluLeuMetAspMetGlnThrPhelYsAla 30
Db 378 GAATTGATGGACATGCAACITTTCAAGCA 407

RESULT 6

ABV28487
ID ABV28487 standard; cDNA; 444 BP.

XX
AC ABV28487;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 28478.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

OS WO200160860-A2.

PN 23-AUG-2001.

PD 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

DR

XX

PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5949; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 444 BP; 144 A; 90 C; 115 G; 93 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.: 3.69e-15 Length: 444
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV28487 (1-444)

Qy 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrIysAlaLeu 20
Db 213 AAGATGGTCTCTCGAGGCTAACTGCGGAATGAAGTTCTATTCCAAACGCTTA 272

Qy 21 GluLeuMetAspMetGlnThrPhelYsAla 30
Db 273 GAATTGATGGACATGCAACITTTCAAGCA 302

RESULT 7

ABV34378
ID ABV34378 standard; cDNA; 444 BP.

XX
AC ABV34378;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 34369.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

OS WO200160860-A2.

PN 23-AUG-2001.

PD 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

DR

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 7224; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 444 BP; 144 A; 91 C; 115 G; 94 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.69e-15 Length: 444
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV34378 (1-444)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db 213 AAAGATGGTCTTCTGAAGGCTAACTCGGATGAAGTTCCTATTCCTCAACTAAGCCTTA 272

QY 21 GluLeuMetAspMetGlnThrPheLysAla 30
Db 273 GAATTGATGCATGCACAACTTTTCAAGCA 302

RESULT 8
ABV22662
ID ABV22662 standard; cDNA; 444 BP.

XX
AC ABV22662;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 22653.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX

PS Claim 1; Page 3975; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 444 BP; 144 A; 90 C; 115 G; 93 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 3.69e-15 Length: 444
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV22662 (1-444)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db 213 AAAGATGGTCTTCTGAAGGCTAACTCGGATGAAGTTCCTATTCCTCAACTAAGCCTTA 272

QY 21 GluLeuMetAspMetGlnThrPheLysAla 30

Db 273 GAATTGATGCATGCACAACTTTTCAAGCA 302

RESULT 9

ABV43239

ID ABV43239 standard; cDNA; 444 BP.

XX
AC ABV43239;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 43230.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8627; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 444 BP; 144 A; 91 C; 115 G; 94 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.69e-15 Length: 444
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV43239 (1-444)

Qy 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db 213 AAGATGCTCTCTGAAGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCCTTA 272

Qy 21 GluLeuMetAspMetGlnThrPhelLysAla 30
Db 273 GAATGTGATGGACATGCAACTTTCAAGCA 302

RESULT 10
AAL10911
ID AAL10911 standard; cDNA; 454 BP.
XX
AC AAL10911;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 3368.
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US000798.
XX
PR 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.
PT
XX Claim 1; Page 620; 3695pp; English.
PS
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
SQ Sequence 454 BP; 125 A; 91 C; 89 G; 148 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 3.8e-15 Length: 454
Score: 152.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 4 Gaps: 0

US-09-489-079-28 (1-47) x AAL10911 (1-454)

Qy 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db 349 AAGATGCTCTCTGAAGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCCTTA 408

Qy 21 GluLeuMetAspMetGlnThrPhelLysAla 30
Db 409 GAATGTGATGGACATGCAACTTTCAAGCA 438

RESULT 11
ABV13085
ID ABV13085 standard; cDNA; 466 BP.
XX
AC ABV13085;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 13076.
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX

PS Claim 1; Page 2161; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX

SQ Sequence 466 BP; 171 A; 87 C; 104 G; 104 T; 0 U; 0 Other;

Alignment Scores: 3.94e-15 Length: 466

Pred. No.: 152.00 Matches: 30

Score: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 61.79% Indels: 0

DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV13085 (1-466)

QY 1 LysAspGlyLeuLeuLysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeu 20

DB 228 AAGATGGTCTTCTGAGGCTAACTGGGATGAAAGTTCTATTCACAAAGCCTTA 287

QY 21 GluLeuMetAspMetGlnThrPheLysAla 30

DB 288 GAATTGATGACATGCAAACTTTCAAAGCA 317

RESULT 12

ACN81108

ID ACN81108 standard; DNA; 470 BP.

AC ACN81108;

XX 02-DEC-2004 (first entry)

XX Breast cancer related marker, seq id 2258.

XX Cancer; breast; tumour; cytostatic; marker; detection; therapy; ds.

XX Homo sapiens.

XX US2003099974-A1.

XX 29-MAY-2003.

XX 18-JUL-2002; 2002US-00198846.

XX 18-JUL-2001; 2001US-0306220P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2003-787014/74.

XX Novel isolated polypeptide associated with breast cancer, useful for

XX detecting presence of polypeptide in sample, as a marker for breast

XX cancer.

XX Disclosure; SEQ ID NO 2258; 36pp; English.

XX The invention relates to an isolated polypeptide (I) associated with

CC breast cancer which is encoded by a nucleic acid molecule comprising a

CC nucleotide sequence (S1). Further disclosed is an antibody that binds to

CC the polypeptide of the invention. The activity of the polypeptide of the

CC invention may be described as cytostatic. The antibody is useful for

CC detecting the presence of (I) in a sample. Nucleic acid molecules of the

CC invention are useful in the detection of breast tumours. (I) is useful as

CC a marker for breast cancer and in breast cancer therapy. Sequences given

CC in records ACN78851-ACN92934 represent nucleic acid markers associated

CC with breast cancer. Note: The sequence listing does not form part of the

CC specification but may be obtained in electronic format from the USPTO web

CC site at seqdata.uspto.gov/sequence.html?DocID=20030099974

XX

SQ Sequence 470 BP; 127 A; 97 C; 93 G; 149 T; 0 U; 4 Other;

Alignment Scores: 3.98e-15 Length: 470

Pred. No.: 152.00 Matches: 30

Score: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 61.79% Indels: 0

DB: 11 Gaps: 0

US-09-489-079-28 (1-47) x ACN81108 (1-470)

QY 1 LysAspGlyLeuLeuLysAlaAenCysGlyMetLysValSerIleProThrLysAlaLeu 20

DB 364 AAGGATGGTCTTCTGAGGCTAACTGGGATGAAAGTTCTATTCACAAAGCCTTA 423

QY 21 GluLeuMetAspMetGlnThrPheLysAla 30

DB 424 GAATTGATGACATGCAAACTTTCAAAGCA 453

RESULT 13

ABV43067

ID ABV43067 standard; cDNA; 473 BP.

XX AC ABV43067;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 43058.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of

XX prostate cells and correlating with presence of prostate cancer, useful

XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 8599-8600; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 473 BP; 162 A; 91 C; 115 G; 105 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.01e-15 Length: 473
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV43067 (1-473)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
 DB 270 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTCTATTCCCAACTAAAGCCTTA 329
 QY 21 GluLeuMetAspMetGlnThrPhelLysAla 30
 DB 330 GAATTGATGGACATGCAAACTTTCAAGCA 359

RESULT 14
 ABV34206
 ID ABV34206 standard; cDNA; 473 BP.
 XX
 AC ABV34206;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 34197.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX

PS Claim 1; Page 7198; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 473 BP; 162 A; 91 C; 115 G; 105 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4.01e-15 Length: 473
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x ABV34206 (1-473)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
 DB 270 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTCTATTCCCAACTAAAGCCTTA 329
 QY 21 GluLeuMetAspMetGlnThrPhelLysAla 30
 DB 330 GAATTGATGGACATGCAAACTTTCAAGCA 359

RESULT 15
 ABV03916
 ID ABV03916 standard; cDNA; 508 BP.
 XX
 AC ABV03916;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 3907.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX

PS Claim 1; Page 692; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 508 BP; 176 A; 100 C; 120 G; 111 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	4,41e-15	Length:	508
Score:	152.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	61.79%	Indels:	0
DB:	5	Gaps:	0

US-09-489-079-28 (1-47) x ABV03916 (1-508)

Qy	1	LysAspGlyLeuLeuLysAlaAanCysGlyMetLysValserIleProThrLysAlaLeu	20
Db	265	AAAGATGGTCTTCTGAAGGCTAACTCGGATGAAGAAGTTCTATTCCAACTAAGCCTTA	324
Qy	21	GluLeuMetAspMetGlnThrPheLysAla	30
Db	325	GAATTGATGGACATGCARAACTTTCARAGCA	354

Search completed: August 1, 2005, 23:18:16
Job time : 91.0391 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:24:58 ; Search time 711.571 Seconds
(without alignments)
3200.517 Million cell updates/sec

Title: US-09-489-079-28

Perfect score: 246

Sequence: 1 KDGLLKANCMKVSIPFKAL.....FKAGKFCNFNFTLRILKY 47

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09489079/runat_01082005_141250_13297/app_query.fasta_1.1635
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1.1 8858 @runat_01082005_141250_13297 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220.5	89.6	85660	9 AL357148	Human DNA
2	220.5	89.6	343019	2 AL391494	Human sapi
3	172.5	70.1	37716	9 AP006507	Human sapi
4	172.5	70.1	93824	9 AP006565	Human sapi

5	172.5	70.1	119182	9 HSB62L20	AL050302 Homo sapi
6	172.5	70.1	162364	9 AP005213	AP005213 Homo sapi
c 7	172.5	70.1	190690	2 AL627234	AL627234 Homo sapi
c 8	172.5	70.1	340000	9 HS21C003	AL163203 Homo sapi
c 9	168	68.3	73775	9 AP003901	AP003901 Homo sapi
10	166	67.5	127500	9 AP006564	AP006564 Homo sapi
11	152	61.8	403	6 CQ472214	CQ472214 Sequence
12	152	61.8	404	6 CQ481383	CQ481383 Sequence
13	152	61.8	434	6 CQ426143	CQ426143 Sequence
14	152	61.8	444	6 CQ490788	CQ490788 Sequence
15	152	61.8	444	6 CQ496633	CQ496633 Sequence
16	152	61.8	444	6 CQ502529	CQ502529 Sequence
17	152	61.8	444	6 CQ511391	CQ511391 Sequence
18	152	61.8	454	6 CQ418338	CQ418338 Sequence
19	152	61.8	466	6 CQ481209	CQ481209 Sequence
20	152	61.8	473	6 CQ502357	CQ502357 Sequence
21	152	61.8	473	6 CQ511219	CQ511219 Sequence
22	152	61.8	508	6 CQ472041	CQ472041 Sequence
23	152	61.8	555	6 CQ509784	CQ509784 Sequence
24	152	61.8	636	6 AX053794	AX053794 Sequence
25	152	61.8	664	6 AX053793	AX053793 Sequence
26	152	61.8	708	6 AX053785	AX053785 Sequence
27	152	61.8	713	6 AX053787	AX053787 Sequence
28	152	61.8	741	6 AX053804	AX053804 Sequence
29	152	61.8	763	6 CQ426355	CQ426355 Sequence
30	152	61.8	766	6 AX053799	AX053799 Sequence
31	152	61.8	770	6 CQ427927	CQ427927 Sequence
32	152	61.8	771	6 AX053802	AX053802 Sequence
33	152	61.8	776	6 AX053782	AX053782 Sequence
34	152	61.8	814	6 AX053780	AX053780 Sequence
35	152	61.8	833	6 AX053800	AX053800 Sequence
36	152	61.8	904	6 AX053778	AX053778 Sequence
37	152	61.8	1158	6 AX053798	AX053798 Sequence
38	152	61.8	1337	6 AR283455	AR283455 Sequence
39	152	61.8	1337	6 AR344223	AR344223 Sequence
40	152	61.8	1337	6 AR351424	AR351424 Sequence
41	152	61.8	1337	6 AR454004	AR454004 Sequence
42	152	61.8	1337	6 AR561592	AR561592 Sequence
43	152	61.8	1337	6 AX282974	AX282974 Sequence
44	152	61.8	1337	6 AX303147	AX303147 Sequence
45	152	61.8	1448	6 CQ490684	CQ490684 Sequence

ALIGNMENTS

RESULT 1
AL357148
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL357148 85660 bp DNA linear PRI 02-MAY-2003
Human DNA sequence from clone RP11-739D18 on chromosome 10,
complete sequence.
AL357148
AL357148.22 GI:30348856
HTG.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Bird.C.
Direct Submission
Submitted (02-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 2, 2003 this sequence version replaced gi:20338435.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information

on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome 10, constructed by the Sanger Centre Chromosome 10
 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr10>
 RP11-739D18 is from the library RPCI-11.3 constructed by the group
 of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACe3.6

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest, except on the rare
 occasion of the clone being a YAC.

FEATURES

source

Location/Qualifiers
 1..85660
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP11-739D18"
 /clone_lib="RPCI-11.3"

ORIGIN

Alignment Scores:

Pred. No.: 1..51e-21 Length: 85660
 Score: 220.50 Matches: 47
 Percent Similarity: 60.26% Conservative: 0
 Best Local Similarity: 60.26% Mismatches: 31
 Query Match: 89.63% Indels: 1
 DB: 9 Gaps: 1

US-09-489-079-28 (1-47) x AL357148 (1-85660)

QY 1 LysAspGlyLeuLeuLys----- 6
 |||||
 Db 36191 AAAGATGGTCTTCTGAAGGTAATAACTTTTATATTTTATCTTGAGTATTAACTACATAT 36250
 |||||
 QY 7 -----AlaAsnCys 9
 |||||
 Db 36251 TTTATGAAGTATACATTGTATATTAATGTTTCTTTCCAAACCCATTAGGCTAACTGC 36310
 |||||
 QY 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 29
 |||||
 Db 36311 GGAATGAAAGTTTCTTATTCCTCACTAAAGCCTTAGAATTGATGACATGCACAACTTTCAA 36370
 |||||
 QY 30 AlaGlyLysPheCysAsnPheAsnPheThrLeuGluArgArgIleLeuLysTyr 47
 |||||
 Db 36371 GCAGGTAATTTTGTAAATTTTAAATTTTACTCTGGAAGAAGAAATATTAATATAT 36424
 |||||

RESULT 2

AL391494/c 343019 bp DNA linear HTG 02-MAY-2003
 LOCUS Homo sapiens chromosome 10 clone RP13-263N18, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 12 unordered pieces.
 ACCESSION AL391494
 VERSION AL391494.9 GI:30348811
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 343019)

AUTHORS

TITLE
JOURNAL

COMMENT

Sims,S.

Direct Submission

Submitted (01-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On May 2, 2003 this sequence version replaced gi:11139963.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: bB263N18

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 180973 bases at least Q40

Consensus quality: 182878 bases at least Q30

Consensus quality: 183958 bases at least Q20

Insert size: 341919; sum-of-contigs

Insert size: 176215; 9.2% error; agarose-fp

Quality coverage: 2.72x in Q20 bases; sum-of-contigs Quality
 coverage: 5.31x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 11652: contig of 11652 bp in length
 * 11653 11752: gap of 100 bp
 * 11753 20071: contig of 8319 bp in length
 * 20072 20171: gap of 100 bp
 * 20172 50720: contig of 30549 bp in length
 * 50721 50820: gap of 100 bp
 * 50821 57213: contig of 6393 bp in length
 * 57214 57313: gap of 100 bp
 * 57314 93039: contig of 35726 bp in length
 * 93040 93139: gap of 100 bp
 * 93140 111822: contig of 18683 bp in length
 * 111823 111922: gap of 100 bp
 * 111923 124154: contig of 12332 bp in length
 * 124155 126375: contig of 2121 bp in length
 * 126376 126475: gap of 100 bp
 * 126476 283525: contig of 157050 bp in length
 * 283526 283625: gap of 100 bp
 * 283626 291539: contig of 7914 bp in length
 * 291540 291639: gap of 100 bp
 * 291640 312365: contig of 20726 bp in length
 * 312366 312465: gap of 100 bp
 * 312466 343019: contig of 30554 bp in length.

FEATURES

source

Location/Qualifiers
 1..343019
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="10"
 /clone="RP13-263N18"
 /clone_lib="RPCI-13.2"

misc_feature

1..11652
 /note="assembly_fragment:00451
 clone end:SP6
 vector side:left"

misc_feature

11753..20071
 /note="assembly_fragment:01291
 fragment_chain:1"
 20172..50720
 /note="assembly_fragment:01448"

FEATURES

```

source
1. .93824
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p"
/clone="RP11-1157N02_B"

ORIGIN
Alignment Scores:
Pred. No.: 3,228-14 Length: 93824
Score: 172.50 Matches: 36
Percent Similarity: 51.28% Conservative: 4
Best Local Similarity: 46.15% Mismatches: 7
Query Match: 70.12% Indels: 31
DB: 9 Gaps: 1

US-09-489-079-28 (1-47) x AP006565 (1-93824)
Qy 1 LysAspGlyLeuLeuYsAla-----7
Db 19672 AAAGATGGTCTCTGAGGTAACTTTATATTTTATCTTGAATTAATTAACCTAT 19731
Qy 8 -----AsnCys 9
Db 19732 TTTATGAAGTATACATTATATAGTAATTATGTGTTCCAAACCCATTAGCCTACCTGT 19791
Qy 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhelys 29
Db 19792 GGAATGAAATTTCTCTTCCAATAAAGCCTTAGAATTGAAGCAGACAGAACATTCAAA 19851
Qy 30 AlaGlyLysPheCysAsnPheAsnPheThrLeuGluAArgAqIleLeuLysTyr 47
Db 19852 GCAGGTAAATTTTGTAATTTTAAATTTTACTGTGGGAATTAAGACATTAAATAT 19905

RESULT 5
HSB62L20 119182 bp DNA linear PRI 01-DEC-1999
LOCUS Homo sapiens chromosome 21 BAC RGB62L20, complete sequence.
DEFINITION AL050302
ACCESSION AL050302
VERSION AL050302.2 GI:6522975
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 119182)
Nordsiek,G., Conrad,A., Dose,S., Grimm,M., Groot,J., Hornischer,K.,
Loehnert,T.H., Nizetic,D., Scharfe,M., Schoen,O., Yaspo,M.L. and
Bloeker,H.
Direct Submission
Submitted (25-MAY-1999) GBF, Dept. of Genome Analysis, Mascheroder
Weg 1, D-38124 Braunschweig, Germany, E-mail: hornischer@gbf.de,
nordsiek@gbf.de, bloecker@gbf.de
On Dec 3, 1999 this sequence version replaced gi:4894179.
Collaborators:
Center for applied Molecular Biology
School of Pharmacy, University of London
29-39 Brunswick Square, London, WC1N 1AX, UK
and
GBF, Dept. of Genome Analysis
Mascheroder Weg 1, D-38124 Braunschweig, Germany
and
Max Planck Institute for Molecular Genetics
Innestrasse 73, D-14195 Berlin-Dahlem, Germany
All annotations in this database entry are developed by
computational tools. It is therefore not explicitly noted in the
feature lines that evidence is not experimental.
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+ + + + +
+ Analysis and annotation were performed with the automatic +
+ 'first-pass' annotation and submission tool +
+ 'AnnoMitter' (Hornischer & Bloeker). +

FEATURES
source
1..119182
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="RGB62L20"
/complement((1..14)..(75..155))
/note="match: ESTs AA825445 AI018653 AI792286 AI792327"
380..426
/note="match: EST AW001671"
complement(436..565)
/note="83% identity: matches 463..592 of consensus"
/rpt_family="L1"
443..564
/note="86% identity: matches 12..133 of consensus"
/rpt_family="AluSz"
568..710
/note="IR1, 73% complementary to IR1' (931..1069)"
/rpt_type=INVERTED
578..677
/note="86% identity: matches 176..272 of consensus"
/rpt_family="AluSz"
complement(618..682)
/note="87% identity: matches 164..228 of consensus"
/rpt_family="L1"
686..729
/note="homology = 79.5%, counts = 11"
/rpt_family="aata repeat"
/rpt_type=TANDEM
689..708
/note="AAAT repeat"

misc_feature
misc_feature
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
satellite

```

```

+
+ Programs used by 'AnnoMitter':
+ + + + +
+ GeneFinder (Green), Vers. 084
+ Organism: human
+ GenScan (Burge & Karlin), Vers. 1.0
+ Used matrix: vertebrate; Minimum score: 0
+ Grail (Xu et al.), Vers. 1.3
+ Organism: human
+ Mzef (Zhang)
+ Prior probability: 0.04; Overlapping number: 0 > Xpound (Thomas
& Skolnick)
+ Base score cutoff: 0.2; Minimal exon length: 3 bp > 'Repeats':
BLASTN 2.0.9 (Altschul et al.)
+ Database(s): * RepBase (human), released 22-DEC-1995
+ * RepBase (primate), released 22-DEC-1995
+ RepBase (mammal), released 22-DEC-1995
+ Minimum identity: 70 %
+ 'ESTs': BLASTN 2.0.9 (Altschul et al.)
+ Database(s): * emb1 (EST), Vers. 60 (16-SEP-1999)
+ * emb1new (EST), Vers. 60+ (26-NOV-1999)
+ Using sequence with
masked repeats
+ Minimum score: 60; Minimum identity: 90 %;
+ 'GSSs': BLASTN 2.0.9 (Altschul et al.)
+ Database(s): * emb1 (GSS), Vers. 60 (16-SEP-1999)
+ * emb1new (GSS), Vers. 60+ (26-NOV-1999)
+ Using sequence with
masked repeats
+ Minimum score: 60; Minimum identity: 90 %;
+ 'Tandem Repeats': GDE 2.2 option 'tandem'
+ Minimum length 2 bp; Maximum length 20 bp; Score threshold 20
+ Treat N's as mismatches? YES; Allow uniform consensi? NO
+ 'Inverted Repeats': GDE 2.2 option 'inverted'
+ 'Micro Satellites': GDE 2.2 option 'sputnik' (Abajian) > 'CpG
Islands': GDE 2.2 option 'cpg'
+ CpG island region size 100 bp;
+ Minimum GC contents 50 %; Observed/Expected 0.6 > 'STS Scan':
e-PCR (Schuler)
+ Margin: 50; Number of mismatches allowed: 0; Word size: 7
+ STS database: 'dbSTS markers'
+ 'trna Scan': tRNAscan-SE (Lowe & Eddy), Vers. 1.11 Sequence
overlaps with Acc Nrs. AL049911 and AL078475.
FEATURES
source
1..119182
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="RGB62L20"
/complement((1..14)..(75..155))
/note="match: ESTs AA825445 AI018653 AI792286 AI792327"
380..426
/note="match: EST AW001671"
complement(436..565)
/note="83% identity: matches 463..592 of consensus"
/rpt_family="L1"
443..564
/note="86% identity: matches 12..133 of consensus"
/rpt_family="AluSz"
568..710
/note="IR1, 73% complementary to IR1' (931..1069)"
/rpt_type=INVERTED
578..677
/note="86% identity: matches 176..272 of consensus"
/rpt_family="AluSz"
complement(618..682)
/note="87% identity: matches 164..228 of consensus"
/rpt_family="L1"
686..729
/note="homology = 79.5%, counts = 11"
/rpt_family="aata repeat"
/rpt_type=TANDEM
689..708
/note="AAAT repeat"

misc_feature
misc_feature
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
satellite

```

```

exon      803..908
/GRATE="GENSCAN prediction, score = 6.80
/NAI, score = 100%, comment = excellent
MZF prediction, score = 0.958"
repeat_region  931..1089
/notes="IR1", 79% complementary to IR1 (568..710)"
/rpt_type=INVERTED
repeat_region  970..1080
/notes="87% identity: matches 175..286 of consensus"
/rpt_family="L1"
repeat_region  complement(973..1080)
/notes="88% identity: matches 157..265 of consensus"
/rpt_family="AluSq"
repeat_region  complement(1124..1209)
/notes="89% identity: matches 37..122 of consensus"
/rpt_family="AluSq"
exon      1407..1478
/notes="MZF prediction, score = 0.601"
exon      1479..1507
/notes="XPOUND prediction, score = 0.672"
misc_feature  (1885..2161)..(1922..2547)
/notes="match: GSSs AG014201 AG014202 AG014200"
misc_feature  complement((1885..2240)..(1922..2547))
/notes="match: GSSs AG014203 AG014204 AG014207 AG014206
AG014205"
repeat_region  2548..2593
/notes="homology = 95.7%, counts = 23"
/rpt_family="AC repeat"
/rpt_type=TANDEM
satellite      2548..2582
/notes="AC repeat"
misc_feature  (2595..3250)..(2828..3328)
/notes="match: GSSs AG014196 AG014199 AG014198 AG014200
AG014197"
misc_feature  complement((2595..3149)..(2839..3328))
/notes="match: GSSs AG014210 AG014209 AG014208 AG014207"
repeat_region  3339..3585
/notes="90% identity: matches 11..257 of consensus"
/rpt_family="AluSc"
repeat_region  complement(3343..3585)
/notes="8% identity: matches 184..424 of consensus"
/rpt_family="L1"
misc_feature  (3636..3811)..(3916..4117)
/notes="match: GSSs AG014196 AG014195 AG014194"
misc_feature  complement(3636..(3802..4038))
/notes="match: GSSs AG014210 AG014211"
exon      complement(4600..4610)
repeat_region  /notes="XPOUND prediction, score = 0.222"
4766..5038
/notes="87% identity: matches 1..278 of consensus"
/rpt_family="AluSq"
repeat_region  complement(4780..5040)
/notes="85% identity: matches 164..424 of consensus"
/rpt_family="L1"
exon      4828..4864
/notes="XPOUND prediction, score = 0.495"
exon      5121..5129
/notes="XPOUND prediction, score = 0.201"
misc_feature  complement(5160..5197)
/notes="match: ESTs AW128901 AW128885"
misc_feature  complement(5164..5201)
/notes="match: GSSs AQ883466"
repeat_region  5251..5286
/notes="homology = 97.2%, counts = 18"
/rpt_family="tg repeat"
/rpt_type=TANDEM
satellite      5255..5286
/notes="tg repeat"
misc_feature  complement((5355..5370)..(5388..5427))
/notes="match: ESTs AI990485 AI356086 AW024170"
exon      5585..5646
/notes="XPOUND prediction, score = 0.256"
misc_feature  (5710..5716)..(5753..5757)

```

```

/note="match: GSSs AG020269 AQ883588"
6036..6283
/note="IR2, 84% complementary to IR2' (7549..7799)"
/rpt_type=INVERTED
repeat_region  6039..6283
/notes="87% identity: matches 9..255 of consensus"
/rpt_family="AluSq"
repeat_region  complement(6045..6170)
/notes="90% identity: matches 301..424 of consensus"
/rpt_family="L1"
misc_feature  complement((6408..6471)..(6669..7116))
/notes="match: GSSs AG013829 AG013831 AG013833 AG013830"
(6546..6741)..7116
misc_feature  /notes="match: GSSs AG013851 AG013852"
complement(6596..6953)
exon      /note="GRAIL, score = 56%, comment = good"
6661..6729
/note="XPOUND prediction, score = 0.372"

```

```

Alignment Scores:
Pred. No.: 4,14e-14 Length: 119182
Score: 172.50 Matches: 36
Percent Similarity: 51.28% Conservative: 4
Best Local Similarity: 46.15% Mismatches: 7
Query Match: 70.12% Indels: 31
DB: 9 Gaps: 1

US-09-489-079-28 (1-47) x HSB62L20 (1-119182)

QY 1 LysApGlyLeuLeuYsala-----7
|||||
Db 70066 AAGATGCTCTTCTGAAGGTAATACTTTTATATTTTATCTTGAATATTAACACTTAT 70125
QY 8 -----AsnCys 9
|||||

Db 70126 TTTATGAAGTATACATATATATAGTAATATTGTTCTTCAACCCATTAGCTACTGT 70185
QY 10 GlyMetLysValserIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 29
|||||
Db 70186 GGAATGAAATTTCTCTCCAAATAAAGCCTTAGAATTGAAGGACAGAGAAACATTCAAA 70245
QY 30 AlaGlyLysPheCysAsnPhenPhenThrLeuGluArgArgIleLeuYsTYT 47
|||||
Db 70246 GCAGGTAAATTTTGTAAATTTTAAATTTTACTGTGTAATTAAGAACATTAATAATAT 70299

RESULT 6
AP005213 Homo sapiens 162364 bp DNA linear PRI 11-JUL-2002
LOCUS sequence.
DEFINITION AP005213
ACCESSION AP005213
VERSION AP005213.3 GI:21728163
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 162364)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2002) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jul 10, 2002 this sequence version replaced gi:21327928.
Location/Qualifiers
COMMENT
FEATURES

```

source

```

1. .162364
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18p"
/clone="RP11-97024"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 5,748-14 Length: 162364
Score: 172.50 Matches: 36
Percent: 51.28% Conservative: 4
Best Local Similarity: 46.15% Mismatches: 7
Query Match: 70.12% Indels: 31
DB: 9 Gaps: 1

```

US-09-489-079-28 (1-47) x AP005213 (1-162364)

```

QY 1 LysAspGlyLeuLeuLysAla-----7
Db 131636 AAAGAATGGCTTCTGAAGGTATAACTTTTATATTTTCTGTAATTAATACTACTAT 131695
QY 8 -----AsnCys 9
Db 131696 TTTATCAAGTATACATTATATAGTAAATTATTTGTTTCCAAACCCATTAGCCTACCTGT 131755
QY 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhelys 29
Db 131756 GGAATGAAATTTCTCTCCAAATAAAGCCTTAGAATTTGAAGGACAGAGAAACATTCAAA 131815
QY 30 AlaGlyLysPheCysAsnPheAsnPheThrLeuGluArgArgIleLeuLysTyr 47
Db 131816 GCAGGTAAATTTTGTAATTTTAAATTTTCTGTTGGAATTAAGAACATTAATAT 131869

```

RESULT 7

```

AL627234/c AL627234 190690 bp DNA linear HTG 31-OCT-2001
LOCUS Homo sapiens chromosome 13 clone RP11-53904, WORKING DRAFT
DEFINITION SEQUENCE, 10 unordered pieces.

```

AL627234

AL627234.8 GI:16596759

HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Source Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Burton, J.

Direct Submission

Submitted (30-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerquest@sanger.ac.uk

On Nov 2, 2001 this sequence version replaced gi:16596727.

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BA53904

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 187938 bases at least Q40

Consensus quality: 188282 bases at least Q30

Consensus quality: 188704 bases at least Q20

Insert size: 189790; sum-of-contigs

Insert size: 195444; 3.4% error; agarose-fp

Quality coverage: 10.55x in Q20 bases; sum-of-contigs Quality

coverage: 10.32x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

* 1 26049: contig of 26049 bp in length
* 26149: gap of 100 bp
* 26150 29660: contig of 3511 bp in length
* 29661 29760: gap of 100 bp
* 29761 37939: contig of 8179 bp in length
* 37940 38039: gap of 100 bp
* 38040 42120: contig of 4081 bp in length
* 42121 42220: gap of 100 bp
* 42221 44683: contig of 2463 bp in length
* 44684 47964: contig of 3181 bp in length
* 47965 48064: gap of 100 bp
* 48065 55758: contig of 7594 bp in length
* 55759 129881: contig of 74123 bp in length
* 129882 129981: gap of 100 bp
* 129982 134693: contig of 4712 bp in length
* 134694 134793: gap of 100 bp
* 134794 190690: contig of 55897 bp in length.

```

FEATURES

source

```

1..190690
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="13"
/clone="RP11-53904"
/clone_lib="RPC1-11.2"

```

1..26049

/note="assembly fragment:00507

fragment chain:1"

26150..29660

/note="assembly fragment:01692

fragment chain:1"

29761..37939

/note="assembly fragment:01192

fragment chain:1"

38040..42120

/note="assembly fragment:04448

fragment chain:2"

42221..44683

/note="assembly fragment:01328

fragment chain:2"

44784..47964

/note="assembly fragment:04679

fragment chain:2"

48065..55658

/note="assembly fragment:03781

fragment chain:3"

55759..129881

/note="assembly fragment:04329

fragment chain:3"

129982..134693

/note="assembly fragment:03636"

134794..190690

/note="assembly fragment:04764"

ORIGIN

Alignment Scores:

```

Pred. No.: 6,8e-14 Length: 190690
Score: 172.50 Matches: 36
Percent Similarity: 51.28% Conservative: 4
Best Local Similarity: 46.15% Mismatches: 7
Query Match: 70.12% Indels: 31
DB: 2 Gaps: 1

```

```
US-09-489-079-28 (1-47) x AL627234 (1-190690)
QY 1 LysAspGlyLeuLeuLysAla-----7
Db 33576 AAGATGGTCTCTGAGAGTAATACTTTTATATTTTATCTTGAATATAACTATTAT 33517
QY 8 -----AenCys 9
Db 33516 TTTATGAAGTATACATTATATAGTAATTATTTGTTTCCAAACCCATTAGCTACCTGT 33457
QY 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 29
Db 33456 GGAATGAAATTTCTCTTCCAAATAAGCCTTAGAATTTGAAGACAGAGAAACATTCAAA 33397
QY 30 AlaGlyLysPheCysAenPheAenPheThrLeuGluArgIleLeuLysTyr 47
Db 33396 GCAGGTAAATTTTGTATTTTAAATTTTACTGTGTAATTAAGACATTAAATAT 33343

RESULT 8
HS21C003/c 340000 bp DNA linear PRI 24-MAY-2000
DEFINITION Homo sapiens chromosome 21 segment HS21C003.
ACCESSION AL163203 AP001658 BA000005
VERSION AL163203.2 GI:7717244
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 340000)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Riesselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K., and
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
Yaspo,M.L.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e.mail: sakaki@sc.riken.go.jp
* URL: http://hgp.gsc.riken.go.jp/
and
* Institute of Molecular Biotechnology, Genome Analysis, *
Beutenbergstrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e.mail: shimizu@mb-med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/
and
* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
* Ihnestrasse 73, D-14195 Berlin, Germany,
* e.mail: info@mpi-zoo.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/
Location/Qualifiers
1. 340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
<1. 14276
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="P127M18, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AJ239318"
<1. 103704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="P16C2, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AL049911"
59858. 179140
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="C1762L20"
/clone_lib="CITB HSP BAC library"
/note="Accession No. AL050302"
104370. 179507
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="P879G12"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AP000026"
131197. 238690
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="P29H4"
/clone_lib="RPC11,3-5 PAC library"
/note="Accession No. AL078475"
237516. 340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q11.1"
/clone="R663H15, 3' partial"
/clone_lib="RPC11-11 BAC library"
/note="Accession No. AP001465"
1749. 2035
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type="DISPERSED"
2039. 2119
/note="MER72"
/rpt_family="LTR/MER4-group"
/rpt_type="DISPERSED"
2530. 2854
/note="MER5A"
```

```

/rpt family="DNA/MER1_type"
3429. .3539
/note="f1A)n"
/rpt family="Simple_repeat"
/rpt type=TANDEM
complement (3576. .3916)
/note="MLT2FB"
/rpt family="LTR/Retroviral"
/rpt type=DISPERSED
4466. .4493
/note="(CA)n"
/rpt family="Simple_repeat"
/rpt type=TANDEM
complement (5074. .6557)
/note="MERS2A"
/rpt family="LTR/MER4-group"
/rpt type=DISPERSED
complement (7237. .7538)
/note="AluY"
/rpt family="SINE/Alu"
/rpt type=DISPERSED
7555. .7623
/note="(CA)n"
/rpt family="Simple_repeat"
/rpt type=TANDEM
complement (7624. .7667)
/note="FAM"
/rpt family="SINE/Alu"
/rpt type=DISPERSED
7794. .8077
/note="AluSq"
/rpt family="SINE/Alu"
/rpt type=DISPERSED
8539. .8630
/note="L1MA9"
/rpt family="LINE/L1"
/rpt type=DISPERSED
8631. .8921
/note="AluX"
/rpt family="SINE/Alu"
/rpt type=DISPERSED
8922. .9552
/note="L1MA9"
/rpt family="LINE/L1"
/rpt type=DISPERSED
9908. .10111
/note="FLAM C"
/rpt family="SINE/Alu"
/rpt type=DISPERSED
10917. .11366
/note="L2"
/rpt family="LINE/L2"
/rpt type=DISPERSED
complement (11847. .12141)
/note="AluX"
/rpt family="SINE/Alu"
/rpt type=DISPERSED
12536. .12642
/note="L2"
/rpt family="LINE/L2"
/rpt type=DISPERSED
12657. .12871
/note="MERS6A"
/rpt family="DNA/MER1_type"
/rpt type=DISPERSED
12879. .13001
/note="L2"
/rpt family="LINE/L2"
/rpt type=DISPERSED
13019. .13503
/note="MLT1D"
/rpt family="LTR/MaLR"

```

```

/rpt_type=DISPERSED

Alignment Scores:
Pred. No.: 1.25e-13 Length: 340000
Score: 172.50 Matches: 36
Percent Similarity: 51.28% Conservative: 4
Best Local Similarity: 46.15% Mismatches: 7
Query Match: 70.12% Indels: 31
DB: 9 Gaps: 1

US-09-489-079-28 (1-47) x HS21C003 (1-340000)

QY 1 LysaspGlyLeuLeuLysAla----- 7
|||||
Db 109016 AAGATGCTCTCTGAAGTAATACTTTATATTTTCTTGAATTAATACTTATCTTGAATTAATACTTAT 108957
|||||

QY 8 -----AsnCys 9
|||||

Db 108956 TTTATGAAGTATACATTATATAGTAATATTGTGTTTCCAAACCCATTAGCCTACCTGT 108897
|||||

QY 10 GlyMetLysValSerIleProThrIysAlaLeuGluLeuMetAspMetGlnThrPheLys 29
|||||
Db 108896 GGAATGAAATTTCTCTTCCAAATAAGCCTTAGAATTGAAGGACAGAGAACATTCAAA 108837
|||||

QY 30 AlaGlyLysPheCysAsnPheThrLeuGluArgArgIleLeuLysTyr 47
|||||
Db 108836 GCAGTAAATTTTGTAAATTTTAAATTTTACTGTGGAATTGAACACATTAAATAT 108783
|||||

RESULT 9
AP003901/c 73775 bp DNA linear PRI 15-MAR-2003
LOCUS Homo sapiens genomic DNA, chromosome 21q, clone:RP5-879G12,
complete sequences.
ACCESSION AP003901
VERSION AP003901.1 GI:14646859
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2001)
REFERENCE
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (05-JUL-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

FEATURES
Location/Qualifiers
1. 73775
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="RP5-879G12"

ORIGIN

Alignment Scores:
Pred. No.: 1.21e-13 Length: 73775
Score: 168.00 Matches: 30
Percent Similarity: 87.18% Conservative: 4
Best Local Similarity: 76.92% Mismatches: 5
Query Match: 68.29% Indels: 0
DB: 9 Gaps: 0

US-09-489-079-28 (1-47) x AP003901 (1-73775)

```

QY 9 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 28
 Db 4529 TGTGGAATGAAAATCTCTCCAAATAAGCCTTAGAATTGAAGCAGACAGAAACATTC 4470
 QY 29 LysAlaGlyLysPheCysAsnPhenThrLeuGluArgArgIleLeuLysTyr 47
 Db 4469 AAAGCAGGTAAATTTTGTAAATTTTAAATTTTACTGTGTAATTAAGAACATTAAATAT 4413

RESULT 10
 AP006564 127500 bp DNA linear PRI 13-AUG-2003
 LOCUS Homo sapiens genomic DNA, chromosome 18 clone:RP11-1157N02_A,
 complete sequence.
 ACCESSION AP006564 AP006280
 VERSION AP006564.1 GI:33620417
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA
 JOURNAL Published Only in Database (2003)
 REFERENCE 2 (bases 1 to 127500)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (08-AUG-2003) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel.81-45-503-9111, Fax.81-45-503-9170)
 COMMENT On Aug 12, 2003 this sequence version replaced gi:32400665.
 FEATURES
 source Location/Qualifiers
 1..127500
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosomes="18"
 /map="18p"
 /clone="RP11-1157N02_A"

ORIGIN
 Alignment Scores:
 Pred. No.: 4,32e-13 Length: 127500
 Score: 166.00 Matches: 31
 Percent Similarity: 84.62% Conservative: 2
 Best Local Similarity: 79.49% Mismatches: 6
 Query Match: 67.48% Indels: 0
 DB: 9 Gaps: 0

US-09-489-079-28 (1-47) x AP006564 (1-127500)

QY 9 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 28
 Db 125989 TGTGGAAGGAAAGTCTCTCCAAATAAGCCTTAGAATTGAAGCAGACAGAAACATTC 126048
 QY 29 LysAlaGlyLysPheCysAsnPhenThrLeuGluArgArgIleLeuLysTyr 47
 Db 126049 AAAGCAGGTAAATTTTGTAAATTTTAAATTTTACTGTGTAATTAAGAACATTAAATAT 126105

RESULT 11
 CQ472214 403 bp DNA linear PAT 30-JAN-2004
 LOCUS Sequence 4081 from Patent WO0160860.
 DEFINITION CQ472214
 ACCESSION CQ472214
 VERSION CQ472214.1 GI:41437833
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
 TITLE Genes differentially expressed in human prostate cancer and their
 use
 JOURNAL Patent: WO 0160860-A 4081 23-AUG-2001;
 FEATURES Millennium Predictive Medicine, Inc. (US)
 source Location/Qualifiers
 1..403
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.34e-13 Length: 403
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 6 Gaps: 0

US-09-489-079-28 (1-47) x CQ472214 (1-403)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
 Db 173 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAGTTTCTATTCCAACTAAGCCTTA 232

QY 21 GluLeuMetAspMetGlnThrPheLysAla 30
 Db 233 GAATTGATGGACATCAACTTTCAAAGCA 262

RESULT 12
 CQ481383 404 bp DNA linear PAT 30-JAN-2004
 LOCUS Sequence 13250 from Patent WO0160860.
 ACCESSION CQ481383
 VERSION CQ481383.1 GI:41447002
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Schlegel,R., Endege,W.O. and Monahan,J.E.
 TITLE Genes differentially expressed in human prostate cancer and their
 use
 JOURNAL Patent: WO 0160860-A 13250 23-AUG-2001;
 FEATURES Millennium Predictive Medicine, Inc. (US)
 source Location/Qualifiers
 1..404
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Alignment Scores:
 Pred. No.: 1.34e-13 Length: 404
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 6 Gaps: 0

US-09-489-079-28 (1-47) x CQ481383 (1-404)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
 Db 173 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAGTTTCTATTCCAACTAAGCCTTA 232

QY 21 GluLeuMetAspMetGlnThrPheLysAla 30

```
Db      233 GAATTGATGGACATGCAAACTTTCAAAGCA 262
RESULT 13
CQ426143
LOCUS      CQ426143          434 bp      DNA
DEFINITION Sequence 11177 from Patent WO0151628.
ACCESSION CQ426143
VERSION    CQ426143.1  GI:41378372
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Lillie,J., Xu,Y., Wang,Y. and Steinmann,K.
TITLE       Novel genes, compositions, kits, and methods for identification,
JOURNAL     assessment, prevention, and therapy of breast cancer
            Patent: WO 0151628-A 11177 19-JUL-2001;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES   Location/Qualifiers
            source
            1..434
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      1.45e-13      Length:      434
Score:          152.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    61.79%      Indels:      0
DB:             6      Gaps:      0

US-09-489-079-28 (1-47) x CQ426143 (1-434)
Qy      1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db      318 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCCTTA 377
Qy      21 GluLeuMetAspMetGlnThrPheLysAla 30
Db      378 GAATTGATGGACATGCAAACTTTCAAAGCA 407

RESULT 14
CQ490788
LOCUS      CQ490788          444 bp      DNA
DEFINITION Sequence 22655 from Patent WO0160860.
ACCESSION CQ490788
VERSION    CQ490788.1  GI:41456407
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE       Genes differentially expressed in human prostate cancer and their
JOURNAL     use
            Patent: WO 0160860-A 22655 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES   Location/Qualifiers
            source
            1..444
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      1.48e-13      Length:      444
Score:          152.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6      Gaps:      0

US-09-489-079-28 (1-47) x CQ490788 (1-444)
Qy      1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db      213 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCCTTA 272
Qy      21 GluLeuMetAspMetGlnThrPheLysAla 30
Db      273 GAATTGATGGACATGCAAACTTTCAAAGCA 302

Query Match:    61.79%      Indels:      0
DB:             6      Gaps:      0

US-09-489-079-28 (1-47) x CQ490788 (1-444)
Qy      1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db      213 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCCTTA 272
Qy      21 GluLeuMetAspMetGlnThrPheLysAla 30
Db      273 GAATTGATGGACATGCAAACTTTCAAAGCA 302

RESULT 15
CQ496633
LOCUS      CQ496633          444 bp      DNA
DEFINITION Sequence 28500 from Patent WO0160860.
ACCESSION CQ496633
VERSION    CQ496633.1  GI:41462252
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1
AUTHORS    Schlegel,R., Endege,W.O. and Monahan,J.E.
TITLE       Genes differentially expressed in human prostate cancer and their
JOURNAL     use
            Patent: WO 0160860-A 28500 23-AUG-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES   Location/Qualifiers
            source
            1..444
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      1.48e-13      Length:      444
Score:          152.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6      Gaps:      0

US-09-489-079-28 (1-47) x CQ496633 (1-444)
Qy      1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db      213 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCCTTA 272
Qy      21 GluLeuMetAspMetGlnThrPheLysAla 30
Db      273 GAATTGATGGACATGCAAACTTTCAAAGCA 302

Search completed: August 2, 2005, 02:53:42
Job time : 851.571 secs
```


Query Match 26.0%; Score 64; DB 2; Length 1441;
Best Local Similarity 33.3%; Pred. No. 2.4;
Matches 13; Conservative 7; Mismatches 15; Indels 4; Gaps 1;

RESULT 8
AB1208
extragenic suppressor protein SubB and to myo-inositol-1(or 4)-monophosphatase homolog 1
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1208
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1208
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <GLA>
A:Cross-references: UNIPROT:O8Y852; GB:NC_003210; PIDN:CAC99144.1; PID:gl6410468; GSPDB:
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo1066
C:Superfamily: suppressor protein subB

Query Match 21.5%; Score 53; DB 2; Length 257;
Best Local Similarity 30.2%; Pred. No. 13;
Matches 16; Conservative 6; Mismatches 15; Indels 16; Gaps 1;
QY 1 KDGLLKANGCMKVS IPT-----KALELMDMTQFKAGKCFN 37
Db 152 KDTLLIANLSVRKPTTMEAVKVSRLHGAASLEYMDVATGRAGVLSAN 204

RESULT 9
F82983
conserved hypothetical protein PA5310 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F82983
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F82983
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-529 <STO>
A:Cross-references: UNIPROT:Q9HTP4; GB:AE004943; GB:AE004915; MID:g9951615; PIDN:AA0869
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA5310
C:Superfamily: cardiolipin synthase

Query Match 21.5%; Score 53; DB 2; Length 529;
Best Local Similarity 40.0%; Pred. No. 28;
Matches 14; Conservative 6; Mismatches 11; Indels 4; Gaps 1;
QY 1 KDGLL-----KANGCMKVS IPTKALELMDMTQFKAG 31
Db 346 KDGLNYLTGKADSGVRRLTNSLEATDVPVAVAG 380

RESULT 10
AI0143
xanthosine phosphorylase (EC 2.4.2.-) [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AI0143
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AI0143
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-287 <KUR>
A:Cross-references: UNIPROT:Q8ZGV4; GB:AL590842; PIDN:CAC90012.1; PID:gl5979234; GSPDB:Q
C:Genetics:
A:Gene: xapA
C:Superfamily: purine-nucleoside phosphorylase
C:Keywords: glycosyltransferase; pentosyltransferase
Query Match 21.3%; Score 52.5; DB 2; Length 287;
Best Local Similarity 37.8%; Pred. No. 17;
Matches 17; Conservative 5; Mismatches 14; Indels 9; Gaps 2;
QY 4 LKANGCMKV-----SIPTKALELMDM-----QTFKAGKCFNFT 39
Db 230 LSAHCGGLKVIALTATNLAEGLSDVLSHEQTLKFAKVASVNF 274

RESULT 11
E96020
probable two-component response regulator protein SMB20722 [imported] - *Sinorhizobium*
C:Species: *Sinorhizobium meliloti*
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: E96020
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E96020
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-225 <KUR>
A:Cross-references: UNIPROT:Q92TS0; GB:AL591985; PIDN:CAC49829.1; PID:gl5141317; GSPDB:Q
A:Experimental source: strain 1021, megaplasmid pSymB
R:Galbert, P.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, P.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
hehault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMB20722
A:Genome: plasmid
C:Superfamily: ompR protein; response regulator homology

Query Match 21.1%; Score 52; DB 2; Length 225;
Best Local Similarity 34.6%; Pred. No. 15;
Matches 9; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
QY 7 ANCGMKVS IPTKALELMDMTQFKAG 32
Db 142 SSAGMPIETIPARELGLELLELFWRAG 167

RESULT 12
T05759
hypothetical protein M4122.160 - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05759
R:Bevan, M.; Reichert, B.J.; Barel, E.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schue
submitted to the Protein Sequence Database, June 1998
A:Reference number: Z15450
A:Accession: T05759

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:04:01 ; Search time 16.7062 Seconds
(without alignments)
1088.086 Million cell updates/sec

Title: US-09-489-079-28
Perfect score: 246
Sequence: 1 KDGLKANGCMKVSIPTRKAL.....PKAGKFCNFTFLRILKY 47

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	246	100.0	47	3 AAB07642	Aab07642 Amino aci
2	152	61.8	168	4 AAB63909	Aab63909 Human pro
3	152	61.8	169	4 AAB63906	Aab63906 Human pro
4	152	61.8	174	4 AAB63926	Aab63926 Human pro
5	152	61.8	185	4 AAB63933	Aab63933 Human pro
6	152	61.8	189	4 AAB63929	Aab63929 Human pro
7	152	61.8	195	4 AAB63918	Aab63918 Human pro
8	152	61.8	207	4 AAB63937	Aab63937 Human pro
9	152	61.8	220	4 AAB63917	Aab63917 Human pro
10	152	61.8	223	4 AAB63903	Aab63903 Human pro
11	152	61.8	225	4 AAB63901	Aab63901 Human pro
12	152	61.8	229	4 AAB63925	Aab63925 Human pro
13	152	61.8	241	7 AAB63983	Aab63983 Human pro
14	152	61.8	266	4 AAB63899	Aab63899 Human pro
15	152	61.8	398	3 AAB07638	Aab07638 Amino aci
16	152	61.8	445	4 AAB50249	Aab50249 Human bre
17	152	61.8	445	4 AAG65987	Aag65987 B726P spl
18	152	61.8	445	4 AAU33350	Aau33350 Human bre
19	152	61.8	445	5 AAB78917	Abj78917 Human bre
20	152	61.8	445	6 ABJ37740	Abj37740 Human tum
21	152	61.8	445	7 ADL93135	Adl93135 Human bre
22	152	61.8	445	8 ADE44425	Ade44425 Human bre
23	152	61.8	650	4 AAB50263	Aab50263 Human bre
24	152	61.8	650	4 AAG65983	Aag65983 B726P spl
25	152	61.8	650	4 AAU33346	Aau33346 Human bre

26	152	61.8	650	5 AAG78913	Abg78913 Human bre
27	152	61.8	650	6 ABJ37736	Abj37736 Human tum
28	152	61.8	650	7 ADL93131	Adl93131 Human bre
29	152	61.8	650	8 ADE44421	Ade44421 Human bre
30	152	61.8	661	6 ABJ37782	Abj37782 Human tum
31	152	61.8	661	7 ADL93214	Adl93214 Human bre
32	152	61.8	743	4 AAU33358	Aau33358 Human bre
33	152	61.8	743	5 AAG78925	Abg78925 Human bre
34	152	61.8	743	6 ABJ37748	Abj37748 Human tum
35	152	61.8	743	7 ADL93156	Adl93156 Human bre
36	152	61.8	1002	4 AAU33351	Aau33351 Human bre
37	152	61.8	1002	5 AAG78918	Abg78918 Human bre
38	152	61.8	1002	6 ABJ37741	Abj37741 Human tum
39	152	61.8	1002	7 ADL93137	Adl93137 Human bre
40	152	61.8	1002	8 ADE44427	Ade44427 Human bre
41	152	61.8	1013	6 ABJ37783	Abj37783 Human tum
42	152	61.8	1013	7 ADL93215	Adl93215 Human bre
43	152	61.8	1095	4 AAU33357	Aau33357 Human bre
44	152	61.8	1095	5 AAG78924	Abg78924 Human bre
45	152	61.8	1095	6 ABJ37747	Abj37747 Human tum

ALIGNMENTS

RESULT 1
AAB07642
ID AAB07642 standard; peptide; 47 AA.

AC AAB07642;

XX 07-NOV-2000 (first entry)

XX Amino acid sequence of an antigenic peptide from BS322 polypeptide.

XX BS322; breast tissue marker; breast disease; breast cancer.

XX Homo sapiens.

XX WO200043420-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000MO-US001452.

XX 21-JAN-1999; 99US-00234716.

XX (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
PI Stroupe SD;

DR WPI; 2000-499217/44.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
PT detection and diagnosis of breast disease especially breast cancer.

XX Claim 23; Page 125; 126pp; English.

PS The present sequence represents an antigenic peptide derived from human
CC BS322 polypeptide. The peptide is used to raise antibodies. BS322 is a
CC breast tissue marker. The BS322 polynucleotides and polypeptides are used
CC to detect and diagnose breast disease, e.g. breast cancer. The BS322
CC polynucleotides are useful as a source of probes and primers, and the
CC BS322 polypeptides are useful as antigens

XX Sequence 47 AA;

Query Match 100.0%; Score 246; DB 3; Length 47;

Best Local Similarity 100.0%; Pred. No. 1.6e-28;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDGLLKANGCMKVSIPTKALELMDMQTFKAGKFCNFTLERRILKY 47
DB 1 KDGLLKANGCMKVSIPTKALELMDMQTFKAGKFCNFTLERRILKY 47

RESULT 2

ID AAB63909 standard; protein; 168 AA.
XX AAB63909;
AC AAB63909;
DT 26-MAR-2001 (first entry)
XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1271.
DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
XX cancer associated antigen; cytostatic; cancer vaccine.
KW Homo sapiens.
XX WO200073801-A2.
OS 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US014749.
XX 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX (LUDW-) LUDWIG INST CANCER RES.
PA Obata Y;
PI WPI; 2001-025274/03.
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.

Example 1; Page 772; 799pp; English.

PS AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
XX represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer

SQ Sequence 168 AA;

Query Match 61.8%; Score 152; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.9e-14; Indels 0; Gaps 0;
Matches 30; Conservative 0; Mismatches 0;

QY 1 KDGLLKANGCMKVSIPTKALELMDMQTFKA 30
DB 63 KDGLLKANGCMKVSIPTKALELMDMQTFKA 92

RESULT 3

ID AAB63906 standard; protein; 169 AA.
XX AAB63906;
AC AAB63906;
XX 26-MAR-2001 (first entry)
DT Human prostate cancer associated antigen protein sequence SEQ ID NO:1268.
XX

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX Homo sapiens.
OS WO200073801-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US014749.
XX 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX (LUDW-) LUDWIG INST CANCER RES.
PA Obata Y;
PI WPI; 2001-025274/03.
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.

Example 1; Page 770-771; 799pp; English.

PS AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
XX represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer

SQ Sequence 169 AA;

Query Match 61.8%; Score 152; DB 4; Length 169;
Best Local Similarity 100.0%; Pred. No. 6e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDGLLKANGCMKVSIPTKALELMDMQTFKA 30
DB 47 KDGLLKANGCMKVSIPTKALELMDMQTFKA 76

RESULT 4
AAB63926 standard; protein; 174 AA.
ID AAB63926
XX AAB63926;
AC AAB63926;

DT 26-MAR-2001 (first entry)

XX Human prostate cancer associated antigen protein sequence SEQ ID NO:1288.
DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX Homo sapiens.

OS WO200073801-A2.
XX 07-DEC-2000.
XX 26-MAY-2000; 2000WO-US014749.
XX 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX Example 1; Page 779; 799pp; English.
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX Sequence 174 AA;
 SQ

Query Match 61.8%; Score 152; DB 4; Length 174;
 Best Local Similarity 100.0%; Pred. No. 6.2e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KDGLLKANGCMKVSIPTKALELMDMOTFKA 30
 DB 108 KDGLLKANGCMKVSIPTKALELMDMOTFKA 137
 |||||

RESULT 5
 AAB63933
 ID AAB63933 standard; protein; 185 AA.
 XX
 AC AAB63933;
 XX 26-MAR-2001 (first entry)
 DT Human prostate cancer associated antigen protein sequence SEQ ID NO:1295.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 OS
 XX WO200073801-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US014749.
 PF
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Obata Y;
 PI
 XX WPI; 2001-025274/03.
 DR Nucleic acids encoding breast, gastric and prostate cancer associated
 XX antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX Example 1; Page 782; 799pp; English.

CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX Sequence 185 AA;
 SQ

Query Match 61.8%; Score 152; DB 4; Length 185;
 Best Local Similarity 100.0%; Pred. No. 6.7e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KDGLLKANGCMKVSIPTKALELMDMOTFKA 30
 DB 46 KDGLLKANGCMKVSIPTKALELMDMOTFKA 75
 |||||

RESULT 6
 AAB63929
 ID AAB63929 standard; protein; 189 AA.
 XX
 AC AAB63929;
 XX 26-MAR-2001 (first entry)
 DT Human prostate cancer associated antigen protein sequence SEQ ID NO:1291.
 DE Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 OS
 XX WO200073801-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 26-MAY-2000; 2000WO-US014749.
 PF
 XX 28-MAY-1999; 99US-0136526P.
 PR 10-SEP-1999; 99US-0153454P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 PA
 XX Obata Y;
 PI
 XX WPI; 2001-025274/03.
 DR Nucleic acids encoding breast, gastric and prostate cancer associated
 XX antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.
 XX Example 1; Page 780; 799pp; English.
 PS
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer
 XX Sequence 189 AA;
 SQ

Query Match 61.8%; Score 152; DB 4; Length 189;
 Best Local Similarity 100.0%; Pred. No. 6.7e-14;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

XX 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Obata Y;
XX WPI; 2001-025274/03.
XX Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
PS Example 1; Page 775; 799pp; English.
XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterized by expression of an abnormal amount of a protein,
CC e.g. cancer.
XX
SQ Sequence 220 AA;
Query Match 61.8%; Score 152; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 8.2e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KDGLLKANGCMKVSIPTKALELMDMQTFKA 30
DB 83 KDGLLKANGCMKVSIPTKALELMDMQTFKA 112
RESULT 10
AAB63903
ID AAB63903 standard; protein; 223 AA.
XX
AC AAB63903;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1265.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.

```

```

XX Example 1; Page 769; 799pp; English.
XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterized by expression of an abnormal amount of a protein,
CC e.g. cancer.
XX
SQ Sequence 223 AA;
Query Match 61.8%; Score 152; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 8.4e-14;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KDGLLKANGCMKVSIPTKALELMDMQTFKA 30
DB 54 KDGLLKANGCMKVSIPTKALELMDMQTFKA 83
RESULT 11
AAB63901
ID AAB63901 standard; protein; 225 AA.
XX
AC AAB63901;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1263.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 99US-0136526P.
PR 10-SEP-1999; 99US-0153454P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer.
XX
PS Example 1; Page 768; 799pp; English.
XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterized by expression of an abnormal amount of a protein,
CC e.g. cancer.
XX

```

```

SQ Sequence 225 AA;
  Query Match          61.8%; Score 152; DB 4; Length 225;
  Best Local Similarity 100.0%; Pred. No. 8.5e-14;
  Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDGLLKANCCKMKSIPPTKALELMDMQTFKA 30
    |||||
Db 91 KDGLLKANCCKMKSIPPTKALELMDMQTFKA 120

RESULT 12
AAB63925
ID AAB63925 standard; protein; 229 AA.
XX
AC AAB63925;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1287.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
  cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
XX
PN WO200073801-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014749.
XX
PR 28-MAY-1999; 98US-0136526P.
PR 10-SEP-1999; 99US-0134549P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
  antigen precursors, useful for diagnosing and treating a condition
  characterized by expression of an abnormal amount of a protein, e.g.
  cancer.
XX
PS Example 1; Page 778-779; 799pp; English.
XX
CC AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
  represent nucleotide sequences encoding human breast, gastric and
  prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63332 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
  represent human breast, gastric and prostate CAAP protein sequence
  respectively. CAAPs have cytostatic activity and can be used in the
  production of cancer vaccines. The human CAAP proteins, peptides, nucleic
  acids or anti-CAAP antibodies are useful for diagnosing and treating a
  condition characterized by expression of an abnormal amount of a protein,
  e.g. cancer
XX
SQ Sequence 229 AA;
  Query Match          61.8%; Score 152; DB 4; Length 229;
  Best Local Similarity 100.0%; Pred. No. 8.7e-14;
  Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDGLLKANCCKMKSIPPTKALELMDMQTFKA 30
    |||||
Db 108 KDGLLKANCCKMKSIPPTKALELMDMQTFKA 137

RESULT 13
ADB83983
ID ADB83983 standard; protein; 241 AA.

```

```

XX ADB83983;
AC
DT 04-DEC-2003 (first entry)
XX
DE Human breast cancer diagnostic marker protein Incyte 411152.
XX
KW cancer; neurodegenerative disorder; human; breast cancer;
  diagnostic marker.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 72 /note= "Encoded by GAG"
FT Misc-difference 75 /note= "Encoded by GAT"
FT Misc-difference 76 /note= "Encoded by AAT"
FT Misc-difference 79 /note= "Encoded by TTT"
FT Misc-difference 83 /note= "Encoded by CCC"
FT Misc-difference 85 /note= "Encoded by AGA"
XX
PN US2003104418-A1.
XX
PD 05-JUN-2003.
XX
PF 25-APR-2002; 2002US-00133757.
XX
PR 27-APR-2001; 2001US-0287153P.
XX
PA (ZHAN/) ZHANG C.
PA (MAHI/) MAHINI B.
PA (WALK/) WALKER M G.
XX
PI Zhang C, Mahini B, Walker MG;
XX
DR WPI; 2003-687833/65.
DR N-PSDB; ADB83986.
XX
PT New combination of polynucleotides, useful for preparing a composition
  for diagnosing or treating cancer or neurodegenerative disorders.
XX
PS Example 11; Page 20-21; 22pp; English.
XX
CC The invention related to a combination of polynucleotides. The
  combination of polynucleotides is useful for preparing a composition for
  diagnosing or treating cancer or neurodegenerative disorders. The present
  sequence represents the amino acid sequence of the human breast cancer
  diagnostic marker Incyte 411152.
XX
SQ Sequence 241 AA;
  Query Match          61.8%; Score 152; DB 7; Length 241;
  Best Local Similarity 100.0%; Pred. No. 9.2e-14;
  Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KDGLLKANCCKMKSIPPTKALELMDMQTFKA 30
    |||||
Db 76 KDGLLKANCCKMKSIPPTKALELMDMQTFKA 105

RESULT 14
AAB63899
ID AAB63899 standard; protein; 266 AA.
XX
AC AAB63899;
XX
DT 26-MAR-2001 (first entry)
XX

```

DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1261.
 XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX

OS Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014749.

XX 28-MAY-1999; 99US-0136526P.

XX 10-SEP-1999; 99US-0153454P.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer.

XX Example 1; Page 767; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytosolic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer

XX SQ Sequence 266 AA;

Query Match 61.8%; Score 152; DB 4; Length 266;
 Best Local Similarity 100.0%; Pred. No. 1e-13;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDGLLKANGCMKVSIPITKALELMDMQTFKA 30
 DB ||||||||||||||||||||||||||||||||
 97 KDGLLKANGCMKVSIPITKALELMDMQTFKA 126

RESULT 15

AAB07638
 ID AAB07638 standard; protein; 398 AA.

XX AC AAB07638;

XX DT 07-NOV-2000 (first entry)

XX Amino acid sequence of BS322 polypeptide.

XX BS322; breast tissue marker; breast disease; breast cancer.

XX OS Homo sapiens.

XX WO200043420-A1.

XX 27-JUL-2000.

XX 21-JAN-2000; 2000WO-US001452.

XX 21-JAN-1999; 99US-00234716.

XX

PA (ABBO) ABBOTT LAB.

XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;

PI Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;

PI Stroupe SD;

XX WPI; 2000-499217/44.

DR N-PSDB; AAA59015.

XX BS322-specific polynucleotides, polypeptides and antibodies, used in the
 PT detection and diagnosis of breast disease especially breast cancer.

XX Claim 23; Page 122-123; 126pp; English.

XX The present sequence represents a human BS322 polypeptide. BS322 is a
 CC breast tissue marker. The BS322 polynucleotides and polypeptides are used
 CC to detect and diagnose breast disease, e.g. breast cancer. The BS322
 CC polynucleotides are useful as a source of probes and primers, and the
 CC BS322 polypeptides are useful as antigens

XX SQ Sequence 398 AA;

Query Match 61.8%; Score 152; DB 3; Length 398;
 Best Local Similarity 100.0%; Pred. No. 1.7e-13;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDGLLKANGCMKVSIPITKALELMDMQTFKA 30
 DB ||||||||||||||||||||||||||||||||
 76 KDGLLKANGCMKVSIPITKALELMDMQTFKA 105

Search completed: August 1, 2005, 22:18:05

Job time : 17.7062 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 1, 2005, 22:05:16 ; Search time 13.1422 Seconds
(without alignments)
1831.334 Million cell updates/sec

Title: US-09-489-079-28

Perfect score: 246

Sequence: 1 KDGLLKANGCMKVS IPTKAL.....FKAGKFCNFTLRLILKY 47

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152	61.8	1341	2 Q9BXX3	Q9bxx3 homo sapien
2	118	48.0	1011	2 Q9BXX2	Q9bxx2 homo sapien
3	66	26.8	1434	2 Q9VF26	Q9vf26 drosophila
4	64.5	26.2	461	2 Q97FR1	Q97fr1 clostridium
5	64	26.0	1441	2 Q26453	Q26453 drosophila
6	60.5	24.6	160	2 Q61J09	Q61j09 drosophila
7	60	24.4	980	2 Q7SGM9	Q7sgm9 neurospora
8	60	24.4	1076	2 Q6FWK3	Q6fwk3 candida gla
9	59	24.0	844	2 Q93552	Q93552 carassius a
10	58	23.6	2483	2 Q70UT1	Q70ut1 cucurbit ye
11	57.5	23.4	1977	2 Q8JUV0	Q8juv0 cucurbit ye
12	57	23.2	481	2 Q726U0	Q726u0 desulfovibr
13	57	23.2	2299	2 Q6A564	Q6a564 bacteroides
14	56.5	23.0	270	2 Q6PUD2	Q6pud2 phoca vitul
15	55.5	22.6	210	2 Q8EJ39	Q8ej39 streptococc
16	55.5	22.6	210	2 Q8E716	Q8e716 streptococc
17	55.5	22.6	268	1 IL1B HORSE	Q28386 equus cabal
18	55.5	22.6	404	1 SAH7 THEME	Q51933 thermotoga
19	55	22.4	236	2 Q94529	Q94529 daucus caro
20	55	22.4	546	2 Q9NW69	Q9nw69 homo sapien
21	55	22.4	577	2 Q9KE87	Q9ke87 bacillus ha
22	55	22.4	1710	2 Q9HIQ1	Q9hiq1 homo sapien
23	55	22.4	1715	2 Q9UFS8	Q9up88 homo sapien
24	54.5	22.2	155	2 Q6CPI8	Q6cpi8 kluyveromyc
25	54	22.0	175	1 V013 RICCN	Q92jq4 rickettsia
26	54	22.0	175	2 Q7PAC1	Q7pac1 rickettsia
27	54	22.0	433	1 SUCB SCHPO	Q94415 schizosacch
28	54	22.0	1252	2 Q96DN3	Q96dn3 homo sapien
29	53.5	21.7	697	2 Q96UC6	Q96uc6 neurospora
30	53.5	21.7	1023	2 Q64Z91	Q64z91 bacteroides
31	53	21.5	257	2 Q8Y852	Q8y852 listeria mo

RESULT 1

Q9BXX3 PRELIMINARY; PRT; 1341 AA.
AC Q9BXX3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Breast cancer antigen NY-BR-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21174979; PubMed=11280766;
RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
Jaeger E., Knuth A., Old L.J., Chen Y.T.;
RT Identification of a Tissue-specific Putative Transcription Factor in
RT Breast Tissue by Serological Screening of a Breast Cancer Library."
RL Cancer Res. 61:2055-2061(2001).
DR EMBL; AF269087; AAK27325.1; -
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0005515; F:protein binding; NAS.
DR GO; GO:0003700; P:transcription factor activity; NAS.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001969; Pept_Asp_AS.
DR Pfam; PF00023; Ank; 6.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 6.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
KW ANK repeat.
SQ SEQUENCE 1341 AA; 152776 MW; 33E53DDE6FD3A58B CRC64;

Query Match 61.8%; Score 152; DB 2; Length 1341;

Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KDGLLKANGCMKVS IPTKALELMDMTFKA 30
667 KDGLLKANGCMKVS IPTKALELMDMTFKA 696

RESULT 2

Q9BXX2 PRELIMINARY; PRT; 1011 AA.
AC Q9BXX2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Breast cancer antigen NY-BR-1.1 (Fragment).
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]_TaxId=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21174979; PubMed=11280766;
 RA Jaeger D., Stockert E., Guere A.O., Scanlan M.J., Karbach J.,
 RA Jaeger E., Knuth A., Old L.J., Chen Y.T.;
 RT "Identification of a Tissue-specific Putative Transcription Factor in
 RT Breast Tissue by Serological Screening of a Breast Cancer Library.";
 RL Cancer Res. 61:2055-2061(2001).
 DR EMBL; AF269088; AAK27326.1; -;
 DR HSP; O75832; UOH.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; ANK; 3.
 DR SMART; SM00248; ANK; 3.
 DR PROSITE; PS50088; ANK_REPEAT; 1.
 DR PROSITE; PS50297; ANK_REPEAT; 1.
 KW ANK repeat.
 FT NON TER 1011 1011
 SQ SEQUENCE 1011 AA; 114250 MW; 16627D7B218DE438 CRC64;
 Query Match 48.0%; Score 118; DB 2; Length 1011;
 Best Local Similarity 73.3%; Pred. No. 9.6e-08;
 Matches 22; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 KDGLKNCGMKVSIPTRKALELMDMOTFKA 30
 DB 613 KDGLKPTCGMKISLPNFKALELKRETFKA 642
 RESULT 3
 QSVF26
 ID Q9VF26 PRELIMINARY; PRT; 1434 AA.
 AC Q9VF26;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG3158-PA.
 GN Name=spn-E; ORFNames=CG3158;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle E.C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abail J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan C., Ferrara S., Fleischmann W.,
 RA Durbin C.J., Evangelista C.C., Ferraz C., Ferrier J., Dunkov B.C., Dunn P.,
 RA Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstein G.W., Weissenbach J.,
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RX FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003710; AAF5235.1; -;
 DR FlyBase; FBgn0003483; spn-E.
 DR GO; GO:0005737; C:cytoplasm; NAS.
 DR GO; GO:0003724; F:RNA helicase activity; TAS.
 DR GO; GO:0006342; P:chromatin silencing; IMP.
 DR GO; GO:0030717; P:karyosome formation; IMP.
 DR GO; GO:0008298; P:mRNA localization, intracellular; IMP.
 DR GO; GO:0007294; P:ocyte cell fate determination (sensu Insecta); IGI.
 DR GO; GO:0001556; P:ocyte maturation; IMP.
 DR GO; GO:0030720; P:ocyte positioning; IMP.
 DR GO; GO:0009949; P:polarity specification of anterior/posterior. . .; IMP.
 DR GO; GO:0009951; P:polarity specification of dorsal/ventral axis; IMP.
 DR GO; GO:0007315; P:pole plasm assembly; NAS.
 DR GO; GO:0007317; P:regulation of pole plasm oskar mRNA localiz. . .; IMP.
 DR GO; GO:0030423; P:RNA interference, targeting of mRNA for des. . .; IMP.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR001650; Helicase_C.

```

DR InterPro; IPR007502; Helicase dom.
DR InterPro; IPR008191; Maternal_tudor.
DR InterPro; IPR002999; Tudor.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00567; TUDOR; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00333; TUDOR; 1.
DR SMART; SM00355; Znf_C2H2; 1.
DR PROSITE; PS00304; TUDOR; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR ATP-binding; Helicase; Hydrolase.
KW SEQUENCE 1434 AA; 164509 MW; 67E8CD39F1484B13 CRC64;

Query Match 26.8%; Score 66; DB 2; Length 1434;
Best Local Similarity 33.3%; Pred. No. 4.4;
Matches 13; Conservative 7; Mismatches 15; Indels 4; Gaps 1;

QY 6 KANCGMKVSIPTKALELMDMTQFKAGKFCNFNFTLERRI 44
DB 139 ETGCGKTTQVPQYILD----EYKSGKYNIVVTQPRRI 173

RESULT 4
Q97FR1 PRELIMINARY; PRT; 461 AA.
AC Q97FR1
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein CAC2666.
GN OrderedLocustNames=CAC2666.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Soucaille P.,
RA Tatubov R.L., Sabathe F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; A8007764; AAK80613.1; -.
DR PIR; B97228; B97228.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 461 AA; 54132 MW; AC320839BE059C03 CRC64;

Query Match 26.2%; Score 64.5; DB 2; Length 461;
Best Local Similarity 45.7%; Pred. No. 2;
Matches 16; Conservative 6; Mismatches 6; Indels 7; Gaps 2;

QY 1 KQGLLKANCGMKVSIPTKALELM-----DMQTFKA 30
DB 31 KDIIIFNSNCIGK--IPTKDLKLAQSSLINDINTFKS 63

RESULT 5
Q26453 PRELIMINARY; PRT; 1441 AA.
AC Q26453
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hle.

```

```

GN Name=hle;
OS Drosophila sp. (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7242;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S; TISSUE=Ovaries;
RX MEDLINE=96033800; PubMed=7590230;
RA Gillespie D.E., Berg C.A.;
RT "Honeleas is required for RNA localization in Drosophila oogenesis and
encodes a new member of the DE-H family of RNA-dependent ATPases.";
RL Genes Dev. 9:2495-2508(1995).
DR EMBL; S79915; AAB35476.2; -.
DR PIR; T13889; T13889.
DR FlyBase; FBgn003483; spn-E.
DR GO; GO:0005737; C:cytoplasm; NAS.
DR GO; GO:0003724; P:RNA helicase activity; TAS.
DR GO; GO:0006342; P:chromatin silencing; IMP.
DR GO; GO:0030717; P:karyosome formation; IMP.
DR GO; GO:0008298; P:RNA localization, intracellular; IMP.
DR GO; GO:0007294; P:oocyte cell fate determination (sensu Insecta); IGI.
DR GO; GO:0001556; P:oocyte maturation; IMP.
DR GO; GO:0030720; P:oocyte positioning; IMP.
DR GO; GO:0009949; P:polarity specification of anterior/posterior. .; IMP.
DR GO; GO:0009951; P:polarity specification of dorsal/ventral axis; IMP.
DR GO; GO:0007315; P:pole plasm assembly; NAS.
DR GO; GO:0007317; P:regulation of pole plasm oskar mRNA localiz. .; IMP.
DR GO; GO:0030423; P:RNA interference, targeting of mRNA for des. .; IMP.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase dom.
DR InterPro; IPR008191; Maternal_tudor.
DR InterPro; IPR002999; Tudor.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF04408; HA2; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00567; TUDOR; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR SMART; SM00333; TUDOR; 1.
DR PROSITE; PS00304; TUDOR; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1441 AA; 164980 MW; 5B52C1666C99347E CRC64;

Query Match 26.0%; Score 64; DB 2; Length 1441;
Best Local Similarity 33.3%; Pred. No. 8.6;
Matches 13; Conservative 7; Mismatches 15; Indels 4; Gaps 1;

QY 6 KANCGMKVSIPTKALELMDMTQFKAGKFCNFNFTLERRI 44
DB 139 ETGCGKTTQVPQYILD----EGYKSGKYNIVVTQPRRI 173

RESULT 6
Q6IJ09 PRELIMINARY; PRT; 160 AA.
AC Q6IJ09
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE HDC16210.
GN ORFNames=HDC16210;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;

```

```

RA Hild M., Beckmann B., Haas S., Koch B., Solovyev V., Busold C.,
RT Fellenberg K., Boutros M., Vingron M., Sauer F., Hohelsel J., Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
RT towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL: BK002907; DAA04112.1; -.
SQ SEQUENCE 160 AA; 18064 MW; E7D5FFA10221CA36 CRC64;

Query Match 24.4%; Score 60.5; DB 2; Length 160;
Best Local Similarity 29.2%; Pred. No. 2.3;
Matches 14; Conservative 9; Mismatches 14; Indels 11; Gaps 1;

Qy 1 KDGLKANGMKVSPPTKALELMDMQTFKAGK-----FCNEN 37
Db 100 RDGYIRRYCGVMNSPSKCGSLTYTFCAAKRTACIRAFRIYFCSFN 147

RESULT 7
Q7SGM9 PRELIMINARY; PRT; 980 AA.
AC Q7SGM9;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE Hypothetical protein.
GN Name=NCU08054.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehm B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Inakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysyellis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Kryatova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Oemani S.A.,
RA DeSouza C.C., Glaes L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AABX01000009; EAA35949.1; -.
DR GO: GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR001764; Glyco_hydro_3N.
DR PRINTS: PR00133; GLHYDRLASE3.
KW PROSITE; PS00775; GLYCOSYL_HYDROL_F3; 1.
KW Hypothetical protein.
SQ SEQUENCE 980 AA; 106688 MW; BEED0C9AB17C7048 CRC64;

Query Match 24.4%; Score 60; DB 2; Length 980;
Best Local Similarity 39.0%; Pred. No. 21;
Matches 16; Conservative 2; Mismatches 23; Indels 0; Gaps 0;

Qy 7 ANCGMKVSPPTKALELMDMQTFKAGKFCNFTLERRILKY 47
Db 907 ANPGQVDFPVKVLGRGFDKVLLEAGKSAKVEFNLRDLISY 947

RESULT 8

```

```

Q6FWK3
ID Q6FWK3 PRELIMINARY; PRT; 1076 AA.
AC Q6FWK3;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Similar to sp|P27636 Saccharomyces cerevisiae YAR019c CDC15.
GN ORFNames=CAGL0C05005g;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284593;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Cattolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekai F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: CR380949; CAG58297.1; -.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO: GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO: GO:0016740; P:transferase activity; IEA.
DR GO: GO:0006458; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000719; Prot_Kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; TyrKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SEQUENCE 1076 AA; 122681 MW; B3F1FE037A16B126 CRC64;

Query Match 24.4%; Score 60; DB 2; Length 1076;
Best Local Similarity 32.0%; Pred. No. 23;
Matches 16; Conservative 11; Mismatches 11; Indels 12; Gaps 2;

Qy 2 DGLLKA-----NCGMKVSPPTKALELMDMQTFKAGKFCNFTFILE 41
Db 514 DDLIRAGIIPVVDYKNGSLVLSIINKHMLMSIQSW--CKWCNPNLPID 561

RESULT 9
Q93552 PRELIMINARY; PRT; 844 AA.
ID Q93552
AC Q93552;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative odorant receptor.
GN Name=GFB1;

```


OS *Carassius auratus* (Goldfish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Olfactory epithelium;
 RX MEDLINE=98426265; PubMed=9751777; DOI=10.1073/pnas.95.20.11987;
 RA Cao Y., Oh B.C., Stryer L.;
 RT "Cloning and localization of two multigene receptor families in
 RT goldfish olfactory epithelium.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:11987-11992 (1998).
 DR EMBL; AF083080; AAC64075.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008067; P:metabotropic glutamate, GABA-B-like recepto. . .; IEA.
 DR GO; GO:0004872; P:receptor activity; IEA.
 DR GO; GO:0007216; P:metabotropic glutamate receptor signaling p. . .; IEA.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR InterPro; IPR011500; NCD3G_GPCR.
 DR InterPro; IPR004073; Vmron_receptor2.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR Pfam; PF07562; NCD3G; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PRINTS; PR01535; VOMERONASL2R.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN 1.
 DR PROSITE; PS00981; G_PROTEIN_REC_F3_3; 1.
 DR PROSITE; PS0259; G_PROTEIN_REC_F3_4; 1.
 DR Receptor.
 KW Receptor.
 SQ SEQUENCE 844 AA; 94598 MW; B0F13B4E7BAC4088 CRC64;
 Query Match 24.0%; Score 59; DB 2; Length 844;
 Best Local Similarity 40.7%; Pred. No. 25;
 Matches 11; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
 QY 9 CGMKVSIPTKALELMDMTQFKAGKFCN 35
 Db 109 CGSRLSTWTATMGLMNSQDFGPGNVCN 135
 RESULT 10
 Q70UT1 PRELIMINARY; PRT; 2483 AA.
 AC Q70UT1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Polypeptide.
 GN Name=p227;
 OS Cucurbit yellow stunting disorder virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 OC Crinivirus.
 OX NCBI_TaxID=51330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Spanish;
 RX MEDLINE=22897980; PubMed=14551825;
 RA Coutts R.H.A., Livieratos I.C.;
 RT "Nucleotide sequence and genome organisation of Cucurbit yellow
 RT stunting disorder virus RNA";
 RL Arch. Virol. 148:2055-2062 (2003).
 DR EMBL; AJ537493; CAD61026.2; -;
 DR GO; GO:0008174; F:mrna methyltransferase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:Viral genome replication; IEA.
 DR InterPro; IPR002160; Prot_inh_Kunz-1g.

DR InterPro; IPR001788; RNA_dep_RNAPol2.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_Psvir.
 DR InterPro; IPR006066; Viral_helicase1.
 DR InterPro; IPR002588; V_methyltrans.
 DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01660; Vmethyltransf; 1.
 DR PROSITE; PS00283; SOYBEAN_KUNITZ; UNKNOWN 1.
 KW Polypeptide.
 FT CHAIN 463 1979 methyltransferase/helicase-like protein.
 FT CHAIN 1980 2483 RNA-dependant RNA polymerase.
 FT CHAIN 1 462 putative leader proteinase.
 SQ SEQUENCE 2483 AA; 286168 MW; 0F56A44058407C84 CRC64;
 Query Match 23.6%; Score 58; DB 2; Length 2483;
 Best Local Similarity 31.9%; Pred. No. 1.2e+02;
 Matches 15; Conservative 5; Mismatches 9; Indels 18; Gaps 1;
 QY 8 NCGMKVSIPTKALELMDMTQF-----KAGKFCNF 36
 Db 1000 NLGGGSASPTPSLELMYITSRYFYRIYVLSWFCNDKNCCKGKFCNF 1046
 RESULT 11
 Q8JUVO PRELIMINARY; PRT; 1977 AA.
 AC Q8JUVO;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE 1a.
 OS Cucurbit yellow stunting disorder virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
 OC Crinivirus.
 OX NCBI_TaxID=51330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22799014; PubMed=12917477; DOI=10.1099/vir.0.19209-0;
 RA Aguilar J.M., Franco M., Marco C.F., Berdiales B.,
 RA Rodriguez-Cerezo B., Truniger V., Aranda M.A.;
 RT "Further variability within the genus Crinivirus, as revealed by
 RT determination of the complete RNA genome sequence of Cucurbit yellow
 RT stunting disorder virus.";
 RL J. Gen. Virol. 84:2555-2564 (2003).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Marco C.F., Aguilar J.M., Abad J., Gomez-Guillamon M.L., Aranda M.A.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY242077; AAM73638.2; -;
 DR GO; GO:0008174; F:mrna methyltransferase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003724; F:RNA helicase activity; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006396; P:RNA processing; IEA.
 DR GO; GO:0019079; P:Viral genome replication; IEA.
 DR InterPro; IPR002160; Prot_inh_Kunz-1g.
 DR InterPro; IPR000606; Viral_helicase1.
 DR InterPro; IPR002588; V_methyltrans.
 DR Pfam; PF01443; Viral_helicase1; 1.
 DR Pfam; PF01660; Vmethyltransf; 1.
 DR PROSITE; PS00283; SOYBEAN_KUNITZ; UNKNOWN 1.
 SQ SEQUENCE 1977 AA; 226812 MW; CFD290023EA10CDB CRC64;
 Query Match 23.4%; Score 57.5; DB 2; Length 1977;
 Best Local Similarity 28.1%; Pred. No. 1.1e+02;
 Matches 16; Conservative 8; Mismatches 10; Indels 23; Gaps 2;
 QY 14 SIPTPSLELMYITSRYFYRIYVLSWFCNDKNCCKGKFCNF-----NFTLERRILKY 47
 Db 1005 SSPTPSLELMYITSRYFYRIYVLSWFCNDKNCCKGKFCNFCAINGWLRKGLLTF 1061

```
RESULT 12
Q726UO
AC Q726UO PRELIMINARY; PRT; 481 AA.
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE B12 binding domain protein/radical SAM domain protein.
GN OrderedLocustNames=DVU3016;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.P., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Dougherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough."
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017319; AAS97487.1; -.
DR TIGR; DVU3016; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR GO; GO:0050897; F: cobalt ion binding; IEA.
DR GO; GO:0005506; F: iron ion binding; IEA.
DR InterPro; IPR006158; B12-binding.
DR InterPro; IPR011006; CheY-like.
DR InterPro; IPR006638; Elp3/MiAe/NifB.
DR InterPro; IPR007197; Radical SAM.
DR Pfam; PF02310; B12-binding; I.
DR Pfam; PF04055; Radical SAM; 1.
DR SMART; SM00723; Elp3; 1.
KW Complete proteome.
SQ SEQUENCE 481 AA; 54075 MW; 5A4FF907A7CBB263 CRC64;

Query Match 23.2%; Score 57; DB 2; Length 481;
Best Local Similarity 42.9%; Pred. No. 26;
Matches 12; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 1 KDGLKANCCKMKSIPITKALELMDQTF 28
|||: ||| ||| ||| ||| |||
Db 371 KDGLPERCGDGDPTSTGELSKLTTY 398

RESULT 13
Q6A564
ID Q6A564 PRELIMINARY; PRT; 2299 AA.
AC Q6A564;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Cellulosomal anchoring scaffoldin B precursor.
GN Name=scaB;
OS Bacteroides cellulosolvens.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=35825;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398177; PubMed=10940036;
EX DOI=10.1128/JB.182.17.4915-4925.2000;
RA Ding S.-Y., Bayer E.A., Steiner D., Shoham Y., Lamed R.;
RT "A scaffoldin of the Bacteroides cellulosolvens cellulosome that
contains 11 type II cohesins."
RL J. Bacteriol. 182:4915-4925(2000).
RN [2]

SEQUENCE FROM N.A.
PubMed=14761991; DOI=10.1128/JB.186.4.968-977.2004;
Xu Q., Bayer E.A., Goldman M., Kenig R., Shoham Y., Lamed R.;
RT "Architecture of the Bacteroides cellulosolvens cellulosome:
description of a cell surface-anchoring scaffoldin and a family 48
cellulase."
RL J. Bacteriol. 186:968-977(2004).
RN [3]
RP SEQUENCE FROM N.A.
RA Ding S.-Y., Bayer E.A., Shoham Y., Steiner D., Lamed R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Xu Q., Bayer E.A., Goldman M., Kenig R., Shoham Y., Lamed R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF224509; AAT79550.1; -.
DR GO; GO:0009274; C: cell wall (sensus Bacteria); IEA.
DR GO; GO:0005576; C: extracellular; IEA.
DR GO; GO:0005179; F: hormone activity; IEA.
DR GO; GO:0002272; P: polysaccharide catabolism; IEA.
DR InterPro; IPR008965; Cellul bind.
DR InterPro; IPR002102; Cohesin.
DR InterPro; IPR000663; Natriopeptide.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00963; Cohesin; 10.
DR Pfam; PF00395; SLH; 3.
DR SMART; SM00183; NAT_PEP; 6.
KW Signal.
FT SIGNAL 1 29 Potential.
FT CHAIN 30 2299 cellulosomal anchoring scaffoldin B.
SQ SEQUENCE 2299 AA; 243625 MW; CAB441B67D4BFC2C CRC64;

Query Match 23.2%; Score 57; DB 2; Length 2299;
Best Local Similarity 28.9%; Pred. No. 1.5e+02;
Matches 13; Conservative 8; Mismatches 20; Indels 4; Gaps 2;

QY 3 GLLKANCCKMKSIPITKALELMDQTFKAGKFCNFNFTLERRILKY 47
|||: ||| ||| ||| ||| ||| ||| |||
Db 502 GKVKVKGADKIKVP---VEIKDIPISIGINN-CNFTLKYSNVLYKY 542

RESULT 14
Q6PUD2
ID Q6PUD2 PRELIMINARY; PRT; 270 AA.
AC Q6PUD2;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Interleukin-1 beta.
OS Phoca vitulina richardsi.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Pinnipedia; Phocidae; Phoca.
OX NCBI_TaxID=271025;
RN [1]
RP SEQUENCE FROM N.A.
RA Bozza M., Atkinson S.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Produced by activated macrophages. IL-1 stimulates
CC thymocyte proliferation by inducing IL-2 release, B-cell
CC maturation and proliferation, and fibroblast growth factor
CC activity (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SIMILARITY: Belongs to the IL-1 family.
DR EMBL; AY578791; AAS91558.1; -.
DR GO; GO:0005576; C: extracellular; IEA.
DR GO; GO:0005149; F: interleukin-1 receptor binding; IEA.
DR GO; GO:0008283; P: cell proliferation; IEA.
DR GO; GO:0006955; P: immune response; IEA.
DR GO; GO:0006954; P: inflammatory response; IEA.
DR GO; GO:0000074; P: regulation of cell cycle; IEA.
DR InterPro; IPR008996; Cytok_IL1_like.
DR InterPro; IPR002348; IL1_HBGF.
DR InterPro; IPR003502; IL1_propep.
```

Search completed: August 1, 2005, 22:22:27
Job time : 18.1422 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 1, 2005, 22:44:18 ; Search time 1787.61 Seconds
(without alignments)
1000.786 Million cell updates/sec

Title: US-09-489-079-28
Perfect score: 246
Sequence: 1 KDGLKANGCMKVSPTKAL.....FKAGKFCNFTLRLIKY 47

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=CGN2_1/USPTO.spool_p/US09489079/runat_01082005_141250_13309/app_query.fasta_1.1635
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09489079 @CGN 1 1 10973 @runat_01082005_141250_13309 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DBLEXT=7

Database :

EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_hic: *
5: gb_est3: *
6: gb_est4: *
7: gb_est5: *
8: gb_est6: *
9: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	220.5	89.6	532	8 B17512	B17512_34502.TPB C
2	220.5	89.6	665	8 B15703	B15703_34502.TP CI
3	168	68.3	685	9 AG010131	AG010131 Homo sapi
4	168	68.3	686	9 AG010132	AG010132 Homo sapi
5	166	67.5	477	8 AQ667184	AQ667184 HS 2109_A
6	161.5	65.7	695	5 AG010124	AG010124 Homo sapi
7	152	61.8	514	5 BP312630	BP312630 BP312630
8	152	61.8	581	5 BP314867	BP314867 BP314867
9	152	61.8	582	5 BP313026	BP313026 BP313026

10	152	61.8	582	5	BP313235	BP313235
11	152	61.8	582	5	BP313436	BP313436
12	152	61.8	582	5	BP314260	BP314260
13	152	61.8	582	5	BP315089	BP315089
14	152	61.8	582	5	BP315806	BP315806
15	152	61.8	582	5	BP312890	BP312890
16	152	61.8	582	5	BP313800	BP313800
17	152	61.8	582	5	BP313822	BP313822
18	152	61.8	582	5	BP314652	BP314652
19	152	61.8	582	5	BP312910	BP312910
20	141	57.3	490	1	BP313704	BP313704
21	127	51.6	482	1	AI951118	wx63g05.x
22	127	51.6	574	2	BP328582	RC5-BN019
23	109	44.3	691	9	AG010133	AG010133 Homo sapi
24	109	44.3	3443	3	BC028407	BC028407 Homo sapi
25	103	41.9	113	2	BF171216	BF171216 PCL2243 M
26	101.5	41.3	344	8	AQ253974	HS 2225_A
27	97	39.4	165	9	HSNC18H11	X88328 H.sapiens-D
28	92	37.4	772	8	BZ878522	CH240_215
29	80	32.5	344	8	AZ576125	AST-T32E0
30	72	29.3	264	9	CR817542	CR817542 GROAAA43A
31	72	29.3	650	9	CR843122	CR843122 GROAAA78A
32	71	28.9	553	8	AQ717052	HS 5469_B
33	70	28.5	661	9	CE154401	tigr-gss-
34	69	28.0	476	9	CG879413	ZMWB050
35	69.5	28.3	476	9	CG879413	ZMWB050
36	69	28.0	541	6	CA046749	CA046749 sealrgb52
37	69	28.0	589	5	BX302871	BX302871 sealrgb52
38	69	28.0	617	6	CB485647	CB485647 otchchm0
39	69	28.0	619	5	BX308194	BX308194
40	69	28.0	620	5	BX302872	BX302872
41	69	28.0	673	5	BX085198	BX085198
42	69	28.0	709	5	BX316527	BX316527
43	69	28.0	713	5	BX316528	BX316528
44	69	28.0	724	6	CA364359	CA364359 639250 NC
45	69	28.0	736	6	CA064133	CA064133 sealrgb53

ALIGNMENTS

RESULT 1

B17512

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

B17512 34502.TPB CIR78SKAL Homo sapiens genomic clone A-345002, genomic survey sequence.

B17512 1 GI:2125261

GSS.

Homo sapiens (human)

Homo sapiens

1 (bases 1 to 532)

Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.

Use of a BAC End Sequence Database for Sequence-Ready Map Building

Unpublished (1997)

Other GSSs: 345002.TP 345002.TVB

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..532

/organism="Homo sapiens"

/mol_type="genomic DNA"

FEATURES
source

/db_xref="taxon:9606"
 /clone="A-345002"
 /sex="Female"
 /cell_type="Fibroblast"
 /clone_lib="CIT978SK1"
 /note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
 CalTech Human BAC Library A1"

ORIGIN

Alignment Scores:
 Pred. No.: 1.9e-23 Length: 532
 Score: 220.50 Matches: 47
 Percent Similarity: 60.26% Conservative: 0
 Best Local Similarity: 60.26% Mismatches: 0
 Query Match: 89.63% Indels: 31
 DB: 8 Gaps: 1

US-09-489-079-28 (1-47) x B17512 (1-532)

QY 1 LysAspGlyLeuLeuLys----- 6
 DB 48 AAAGATGGTCTCTGAAGTAATACTTTTATATTTTCTTGAGTAGTAACTACATAT 107
 QY 7 -----AlaAsnCys 9
 DB 108 TTTATGAAGTATACATTGTATATTATTTCTTTCCAAACCCATTTAGGCTAACTGC 167
 QY 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhelys 29
 DB 168 GGAATGAAAGTTTCTATTCCTCAACTAAAGCTTAGAATTGATGGACATGCAAACTTTCAA 227
 QY 30 AlaGlyLysPheCysAsnPheAsnPheThrLeuGluArgArgIleLeuLysTyr 47
 DB 228 GCAGGTAAATTTTGTAATTTTAAATTTTCTCTGGAAGAAGAAATATTAAATAT 281

RESULT 2

B15703
 LOCUS B15703 C17978SK1 Homo sapiens genomic clone A-345002, genomic survey sequence.
 DEFINITION

ACCESSION B15703
 VERSION B15703.1 GI:2123452
 KEYWORDS GSS.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 665)

AUTHORS Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.

TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building

JOURNAL Unpublished (1997)

COMMENT Other_GSSs: 345002.TPB 345002.TVB

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html

Seq primer: S86

Class: BAC ends.

FEATURES

source

Location/Qualifiers
 1..665
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="A-345002"
 /sex="Female"
 /cell_type="Fibroblast"
 /clone_lib="CIT978SK1"

/note="Vector: pBAC108L; Site_1: HindIII; Site_2: HindIII;
 CalTech Human BAC Library A1"

ORIGIN

Alignment Scores:
 Pred. No.: 2.53e-23 Length: 665
 Score: 220.50 Matches: 47
 Percent Similarity: 60.26% Conservative: 0
 Best Local Similarity: 60.26% Mismatches: 0
 Query Match: 89.63% Indels: 31
 DB: 8 Gaps: 1

US-09-489-079-28 (1-47) x B15703 (1-665)

QY 1 LysAspGlyLeuLeuLys----- 6
 DB 34 AAAGATGGTCTCTGAAGTAATACTTTTATATTTTCTTGAGTAGTAACTACATAT 93
 QY 7 -----AlaAsnCys 9
 DB 94 TTTATGAAGTATACATTGTATATTATTTCTTTCCAAACCCATTTAAGCTAACTGC 153
 QY 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhelys 29
 DB 154 GGAATGAAAGTTTCTATTCCTCAACTAAAGCTTAGAATTGATGGACATGCAAACTTTCAA 213
 QY 30 AlaGlyLysPheCysAsnPheAsnPheThrLeuGluArgArgIleLeuLysTyr 47
 DB 214 GCAGGTAAATTTTGTAATTTTAAATTTTCTCTGGAAGAAGAAATATTAAATAT 267

RESULT 3

AG010131
 LOCUS AG010131 Homo sapiens genomic DNA, 21q region, clone: f2G2X4, genomic survey sequence.
 DEFINITION

ACCESSION AG010131 AG003942
 VERSION AG010131.1 GI:3294407
 KEYWORDS GSS.

SOURCE

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 685)

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.

TITLE Homo sapiens genomic DNA, chromosome 21q

JOURNAL Published Only in Database (1998)

REFERENCE 2 (bases 1 to 685)

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.

TITLE Direct Submission

JOURNAL Submitted (08-JUL-1998) Masahira Hattori, Kitasato University,

Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,

Sagamihara 228, Japan (E-mail:hattori@hgc.ims.u-tokyo.ac.jp,

Tel:0427-78-9732, Fax:0427-78-9561)

On Feb 5, 1999 this sequence version replaced gi:2760801.

AG003942: Submitted (09-Jan-1998).

Location/Qualifiers

1..685

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q"

/clone="f2G2X4"

ORIGIN

Alignment Scores:

Pred. No.: 3.25e-15 Length: 685
 Score: 168.00 Matches: 30
 Percent Similarity: 87.18% Conservative: 4
 Best Local Similarity: 76.92% Mismatches: 5
 Query Match: 68.23% Indels: 0
 DB: 9 Gaps: 0

US-09-489-079-28 (1-47) x AG010131 (1-685)

QY 9 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 28
 Db 74 TGTGGATGAAATTTCTCTTCCAAATAAGCTTAGAATTGAAGCAGACAGAAACATTC 133
 QY 29 LysAlaGlyLysPheCysAsnPheAenPheThrLeuGluArgArgIleLeuLysTyr 47
 Db 134 AAAGCAGGTAATTTTGTAAATTTTAAATTTTACTGTGGAATTAAAGACATTAATAATAT 190

RESULT 4
 AG010132 686 bp DNA linear GSS 14-APR-1999
 LOCUS Homo sapiens genomic DNA, 21q region, clone: f2G2X4, genomic survey
 DEFINITION sequence.

ACCESSION AG010132 AG003943
 VERSION AG010132.1 GI:3294408
 KEYWORDS GSS.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 686)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.

JOURNAL Published Only in DataBase (1998)

REFERENCE 2 (bases 1 to 686)
 Hattori, M., Ishii, K., Toyoda, A., Shiba, T. and Sakaki, Y.

AUTHORS Direct Submission

JOURNAL Submitted (08-JUL-1998) Masahira Hattori, Kitasato University,
 Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,
 Sagami-hara 228, Japan (E-mail: hattori@hgc.ims.u-tokyo.ac.jp,
 Tel: 0427-78-9732, Fax: 0427-78-9561)

COMMENT On Feb 5, 1999 this sequence version replaced gi:2760802.
 AG003943: Submitted (09-Jan-1998).

FEATURES
 source
 1..686
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q"
 /clone="f2G2X4"

ORIGIN

Alignment Scores:
 Pred. No.: 3,25e-15 Length: 686
 Score: 168.00 Matches: 30
 Percent Similarity: 87.18% Conservative: 4
 Best Local Similarity: 76.92% Mismatches: 5
 Query Match: 68.29% Indels: 0
 DB: 9 Gaps: 0

US-09-489-079-28 (1-47) x AG010132 (1-686)

QY 9 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 28
 Db 47 TGTGGATGAAATTTCTCTTCCAAATAAGCTTAGAATTGAAGCAGACAGAAACATTC 106
 QY 29 LysAlaGlyLysPheCysAsnPheAenPheThrLeuGluArgArgIleLeuLysTyr 47
 Db 107 AAAGCAGGTAATTTTGTAAATTTTAAATTTTACTGTGGAATTAAAGACATTAATAATAT 163

RESULT 5
 AQ667184 477 bp DNA linear GSS 24-JUN-1999
 LOCUS HS 2109 A1 A08 T7C CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate=2109 Col=15 Row=A, genomic survey
 sequence.

ACCESSION AQ667184
 VERSION AQ667184.1 GI:5199930
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 477)

AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome

JOURNAL proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

PUBMED 10449764

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 2109 row: A column: 15

Seq primer: T7

Class: BAC ends

High quality sequence stop: 477.

FEATURES

source

1..477
 Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="Plate=2109 Col=15 Row=A"

/sex="male"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
 E-Coli DH10B"

ORIGIN

Alignment Scores:
 Pred. No.: 4,17e-15 Length: 477
 Score: 166.00 Matches: 31
 Percent Similarity: 84.62% Conservative: 2
 Best Local Similarity: 79.49% Mismatches: 6
 Query Match: 67.48% Indels: 0
 DB: 8 Gaps: 0

US-09-489-079-28 (1-47) x AQ667184 (1-477)

QY 9 CysGlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPhe 28
 Db 142 TGTGGAGGAAAGTTTCTCTTCCAAATAAGCTTTAGAATTGAAGCAGACAGAAACATTC 201

QY 29 LysAlaGlyLysPheCysAsnPheAenPheThrLeuGluArgArgIleLeuLysTyr 47
 Db 202 AAAGCAGGTAATTTTGTCAATTTTAAATTTTACTCTGGAATTAAAGATATTAACATAT 258

RESULT 6

AG010124/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE Direct Submission
JOURNAL Submitted (08-JUL-1998) Masahira Hattori, Kitasato University, Department of Science, JST Sequencing Laboratory; Kitasato 1-15-1, Sagamihara 228, Japan [E-mail:hattori@hgc.ims.u-tokyo.ac.jp, Tel:0427-78-9732, Fax:0427-78-9561]
COMMENT On Feb 5, 1999 this sequence version replaced gi:2760794.
FEATURES AG003935: Submitted (05-Jan-1998).

source
Location/Qualifiers

1..695
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="21"
 /map="21q"
 /clone="f2G2X4"

ORIGIN

Alignment Scores:
 Pred. No.: 3,32e-14 Length: 695
 Score: 161.50 Matches: 35
 Percent Similarity: 50.00% Conservative: 4
 Best Local Similarity: 44.87% Mismatches: 8
 Query Match: 65.65% Indels: 31
 DB: 9 Gaps: 1

US-09-489-079-28 (1-47) x AG010124 (1-695)

Qy 1 LysAspGlyLeuLeuLysAlaAAsnCysGlyMetLysValSerIleProThrLysAlaLeu 695
 |||||
 391 AAAGATGGTCTCTCGAAGGTAATACTTTTATATTTTCTTGAATTAATACTACTTAT 332
 |||||

Qy 8 -----AsnCys 9

Db 331 TTTATGAAGTATACATTATATAGTAATTTATGTGTTTCCAAACCACTTACCTGT 272
 |||||

Qy 10 GlyMetLysValSerIleProThrLysAlaLeuGluLeuMetAspMetGlnThrPheLys 29
 |||||
 271 GGAATGAATAATTTCTCTTCCAAATAAGCCTTAGAATTGAAGCAGAGAAACATTCAA 212
 |||||

Qy 30 AlaGlyLysPheCysAsnPheThrLeuGluArgGileLeuLysTyr 47
 |||||

Db 211 GCAGGTAATTTNGTAATTTTAAATTTTCTGTGGAATTGAACACATTAATAAT 158
 |||||

RESULT 7
BP312630
LOCUS BP312630 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION CDNA clone OPR00731, mRNA sequence.
ACCESSION BP312630
VERSION BP312630.1 GI:52241605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.

source
Location/Qualifiers

1..514
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OPR00731"
 /tissue_type="mammary gland"

/cell_line="OCUB-F"
 /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 6.6e-13 Length: 514
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP312630 (1-514)

Qy 1 LysAspGlyLeuLeuLysAlaAAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
 |||||
 312 AAAGATGGTCTCTCGAAGGTAATACTGCGAATGAAGTTCTATTCCAACTAAAGCCTTA 371
 |||||

Qy 21 GluLeuMetAspMetGlnThrPheLysAla 30
 |||||

Db 372 GAATTGATGGACATGCAAACTTTTCAGCA 401
 |||||

RESULT 8

BP314867

LOCUS

BP314867 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION CDNA clone OPR07276, mRNA sequence.
ACCESSION BP314867
VERSION BP314867.1 GI:52243842
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 581)
 Suzuki,Y., Yamashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES

source

1..581
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OPR07276"
 /tissue_type="mammary gland"
 /cell_line="OCUB-F"
 /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:
 Pred. No.: 7.71e-13 Length: 581
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP314867 (1-581)

Qy 1 LysAspGlyLeuLeuLysAlaAAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
 |||||
 268 AAAGATGGTCTCTCGAAGGTAATACTGCGAATGAAGTTCTATTCCAACTAAAGCCTTA 327
 |||||

Qy 21 GluLeuMetAspMetGlnThrPheLysAla 30
 |||||


```
|||||
328 GAATTGATGACATGCACAACTTTCAAAGCA 357

DB
RESULT 9
BP313026          582 bp.  mRNA   linear   EST 17-SEP-2004
LOCUS             Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION        cDNA clone OFR01939, mRNA sequence.
ACCESSION         BP313026
VERSION           BP313026.1 GI:52242001
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 582)
AUTHORS          Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE            Sequence comparison of human and mouse genes reveals a homologous
JOURNAL           block structure in the promoter regions
COMMENT          Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR01939"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores: 7.73e-13 Length: 582
Pred. No.: 152.00 Matches: 30
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP313026 (1-582)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
|||||
DB 214 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTA 273
|||||
QY 21 GluLeuMetAspMetGlnThrPheLysAla 30
|||||
DB 274 GAATTGATGACATGCACAACTTTCAAAGCA 303
|||||

ORIGIN
Alignment Scores: 7.73e-13 Length: 582
Pred. No.: 152.00 Matches: 30
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP313026 (1-582)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
|||||
DB 214 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTA 273
|||||
QY 21 GluLeuMetAspMetGlnThrPheLysAla 30
|||||
DB 274 GAATTGATGACATGCACAACTTTCAAAGCA 303
|||||

RESULT 10
BP313235          582 bp.  mRNA   linear   EST 17-SEP-2004
LOCUS             Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION        cDNA clone OFR02569, mRNA sequence.
ACCESSION         BP313235
VERSION           BP313235.1 GI:52242210
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 582)
AUTHORS          Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE            Sequence comparison of human and mouse genes reveals a homologous
JOURNAL           block structure in the promoter regions
COMMENT          Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR02569"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores: 7.73e-13 Length: 582
Pred. No.: 152.00 Matches: 30
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP313235 (1-582)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
|||||
DB 112 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTA 171
|||||
QY 21 GluLeuMetAspMetGlnThrPheLysAla 30
|||||
DB 172 GAATTGATGACATGCACAACTTTCAAAGCA 201
|||||

RESULT 11
BP313436          582 bp.  mRNA   linear   EST 17-SEP-2004
LOCUS             Sugano cDNA library, mammary gland OCUB-F Homo sapiens
DEFINITION        cDNA clone OFR03209, mRNA sequence.
ACCESSION         BP313436
VERSION           BP313436.1 GI:52242411
KEYWORDS          EST.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE         1 (bases 1 to 582)
AUTHORS          Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
Mizushima-Sugano,J., Nakai,K. and Sugano,S.
TITLE            Sequence comparison of human and mouse genes reveals a homologous
JOURNAL           block structure in the promoter regions
COMMENT          Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers

FEATURES
source
1..582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OFR03209"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores: 7.73e-13 Length: 582
Pred. No.: 152.00 Matches: 30
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.79% Indels: 0
DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP313235 (1-582)

QY 1 LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
|||||
DB 112 AAAGATGGTCTTCTGAAGGCTAACTGCGGAATGAAAGTTTCTATTCCAACTAAAGCCTTA 171
|||||
QY 21 GluLeuMetAspMetGlnThrPheLysAla 30
|||||
DB 172 GAATTGATGACATGCACAACTTTCAAAGCA 201
|||||
```

Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP313436 (1-582)

QY 1 LysAspGlyLeuLeuIysAlaAsnCysGlyMetIysValSerIleProThrIysAlaLeu 20
 |||||
 DB 239 AAAGATGGTCTTCTGAAGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCTTA 298
 |||||

QY 21 GluLeuMetAspMetGlnThrPheIysAla 30
 |||||
 DB 299 GAATTGATGGACATGCAAACTTTCAAAGCA 328
 |||||

RESULT 12
 BP314260
 LOCUS BP314260 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION CDNA clone OFR05703, mRNA sequence.

ACCESSION BP314260
 VERSION BP314260.1 GI:52243235
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki

Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES

source

1..582
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="OFR05703"
 /tissue_type="mammary gland"
 /cell_line="OCUB-F"
 /clone_lib="Sugano cDNA library, mammary gland OCUB-F"
 /note="mammary gland tumor"

ORIGIN

Alignment Scores:

Pred. No.: 7,73e-13 Length: 582
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP314260 (1-582)

QY 1 LysAspGlyLeuLeuIysAlaAsnCysGlyMetIysValSerIleProThrIysAlaLeu 20
 |||||
 DB 114 AAAGATGGTCTTCTGAAGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCTTA 173
 |||||

QY 21 GluLeuMetAspMetGlnThrPheIysAla 30
 |||||
 DB 174 GAATTGATGGACATGCAAACTTTCAAAGCA 203
 |||||

RESULT 13

BP315089

LOCUS BP315089 Sugano cDNA library, mammary gland OCUB-F Homo sapiens
 DEFINITION CDNA clone OFR09418, mRNA sequence.

CDNA clone OFR07766, mRNA sequence.

BP315089

VERSION BP315089.1 GI:52244064

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

1 (bases 1 to 582)

Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

1..582

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="OFR07766"

/tissue_type="mammary gland"

/cell_line="OCUB-F"

/clone_lib="Sugano cDNA library, mammary gland OCUB-F"

/note="mammary gland tumor"

ORIGIN

Alignment Scores:

Pred. No.: 7,73e-13 Length: 582
 Score: 152.00 Matches: 30
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 61.79% Indels: 0
 DB: 5 Gaps: 0

US-09-489-079-28 (1-47) x BP315089 (1-582)

QY

1 LysAspGlyLeuLeuIysAlaAsnCysGlyMetIysValSerIleProThrIysAlaLeu 20

|||||

DB 172 AAAGATGGTCTTCTGAAGCTAACTGCGGAATGAAAGTTCTATTCCAACTAAAGCCTTA 231

|||||

QY 21 GluLeuMetAspMetGlnThrPheIysAla 30

|||||

DB 232 GAATTGATGGACATGCAAACTTTCAAAGCA 261

|||||

RESULT 14

BP315806

LOCUS BP315806 Sugano cDNA library, mammary gland OCUB-F Homo sapiens

DEFINITION CDNA clone OFR09418, mRNA sequence.

ACCESSION BP315806

VERSION BP315806.1 GI:52244781

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 582)

Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: yusuzuki@ims.u-tokyo.ac.jp.

Location/Qualifiers

```

Qy      1  LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
      .
Db     128  AAGATGGTCTTCTGAAGCTAACTGCCGAATGAAGTTCTATTCCAACATAAAGCCCTTA 187

Qy     21  GluLeuMetAspMetGlnThrPhelLysAla 30
      .
Db    188  GAATTGATGGACATGCNAACITTCAAAGCA 217

Search completed: August 2, 2005, 11:46:17
Job time : 1791.61 secs

```

```

source
1. .582
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR09418"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.:      7,73e-13      Length:      582
Score:          152.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      61.79%      Indels:      0
DB:              5          Gaps:      0

US-09-489-079-28 (1-47) x BP315806 (1-582)

QY      1  LysAspGlyLeuLeuLysAlaAsnCysGlyMetLysValSerIleProThrLysAlaLeu 20
Db      114  AAGATGGTCTCTGAGGCTAACTGCGGATGAAAGTTCTATTCCAACTAAGCCTTA 173
QY      21  GluLeuMetAspMetGlnThrPheLysAla 30
Db      174  GAATTGATGGACATGCAAACTTTCAAAGCA 203

RESULT 15
BP312890
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. .583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="OPR01356"
/tissue_type="mammary gland"
/cell_line="OCUB-F"
/clone_lib="Sugano cDNA library, mammary gland OCUB-F"
/note="mammary gland tumor"

ORIGIN
Alignment Scores:
Pred. No.:      7,75e-13      Length:      583
Score:          152.00      Matches:      30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      61.79%      Indels:      0
DB:              5          Gaps:      0

US-09-489-079-28 (1-47) x BP312890 (1-583)

```

This Page Blank (uspto)